Role of Co-transcriptional R-loops in DNA Replication Stress

DISSERTATION
ZUR
ERLANGUNG DER NATURWISSENSCHAFTLICHEN DOKTORWÜRDE
(Dr. sc. nat.)

VORGELEGT DER
MATHEMATISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT
DER
UNIVERSITÄT ZÜRICH
VON

SHRUTI MENON

AUS INDIEN

PROMOTIONSKOMMISSION:

PROF. Dr. KERSTIN GARI (Vorsitz)
PD Dr. PAVEL JANSCAK (Leitung der Dissertation)
PROF. Dr. ALESSANDRO SARTORI
PROF. Dr. PRIMO SCHÄR

ZÜRICH, 2018

TABLE OF CONTENTS

ZUSAMMENFASSUNG

Ein Hauptcharakteristikum bei Krebs ist die genomische Instabilität. Sie wird besonders durch Replikationsstress verursacht, speziell durch das verlangsamte oder blockierte Voranschreiten der Replikationsgabeln, was die Entstehung von DNA-Doppelstrangbrüchen begünstigen kann. Es existieren spezifische Loci, sog. "fragile sites", die sich unter Replikationsstress als besonders bruchempfindlich erweisen. Diese Loci treten gehäuft in Genomregionen auf, wo in Krebszellen chromosomale Umgruppierungen auftreten. Onkogenaktivierung in präkanzerösen Läsionen verursacht Replikationsstress, dabei ist ein Hauptauslöser für Stress die Aktivierung stiller Replikationsstartpunkte, wodurch es zu einer gegenseitigen räumlichen Behinderung der Replikations- und Transkriptionskomplexe kommen kann und möglicherweise die Entstehung ko-transkriptioneller R-loops begünstigt wird. R-loops sind drei-strängige Nukleinsäurestrukturen, die entstehen, wenn die wachsende mRNA wieder an den codogenen DNA-Strang bindet, sodass sich ein RNA-DNA-Hybrid und eine Schleife aus einzelsträngiger DNA formen.

Diese Dissertation untersucht die Bedeutung der R-loops unter Bedingungen mit Replikationsstress. Um R-loops zu detektieren, wurde eine Osteosarkom-Zelllinie (U2OS) mit einer katalytisch inaktiven Mutante der humanen Endonuklease RNase H1 entwickelt. Während das Wildtyp-Protein die R-loops behebt, indem es die RNA aus dem Hybrid abspaltet, kann die mutierte RNase H1 lediglich an die Hybridstruktur binden, diese aber nicht abbauen. Durch die am R-loop gebundene RNase H1 wird dieser mittels Fluoreszenz-mikroskopie sichtbar oder kann durch Chromatin-Immunpräzipitation isoliert werden. Bedingungen der Onkogenaktivierung wurden mittels chemischer Substanzen ausgelöst. Der dabei induzierte Replikationsstress korrelierte sehr stark mit der Entstehung der R-loops. Diese R-loops stellen ihrerseits ein zusätzliches Hindernis für die Progression der Replikationsgabel dar. Auch beim Eintritt der Zelle in die Mitose sind persistierende R-loops ursächlich mitbeteiligt an Defekten der Chromosomensegregation.Eine genomweite Analyse zeigte, dass unter Replikationsstress R-loops an bestimmten

Loci entstehen, die signifikant mit chromosomalen fragile sites übereinstimmen. So konnte aufgezeigt werden, dass unter Replikationsstress die Entstehung von R-loops zur genomischen Instabilität beiträgt.

SUMMARY

Genomic instability is a hallmark of cancer. One major cause of this instability is replication stress, slowing or stalling of replication forks, which can lead to DNA double-strand breaks. There are specific loci, termed fragile sites, which are particularly susceptible to breakage under conditions of replication stress. These loci coincide with recurrent sites of chromosomal rearrangements in cancer. Oncogene activation, in pre-cancerous lesions, causes replication stress. A major source of this stress stems from the activation of dormant origins, which can cause conflicts between transcription-replication complexes and possibly formation of co-transcriptional R-loops. R-loops are three stranded nucleic acid structures that form when the nascent mRNA released from a transcribing RNA polymerase reanneals with the template DNA.

In this study, we wanted to address the role of R-loops in replication stress. We first developed a tool to study R-loops in U2OS cells based on a catalytically-inactive mutant of human RNase H1 - an endonuclease that resolves these structures by cleaving the RNA strand in the hybrid. The mutation results in an enzyme that binds to but does not degrade hybrids. This tool enables us to visualize R-loops by fluorescence microscopy and isolate them by chromatin immunoprecipitation (ChIP). Using chemical inducers of replication stress to mimic oncogene activation, we could show that replication stress is associated with formation of R-loops. These R-loops pose a further block to replication fork progression and unresolved R-loops can then cause problems when cells enter mitosis, contributing to defects in chromosome segregation. Genome-wide analysis revealed that R-loops are formed at distinct loci under replication stress and these sites coincide significantly with chromosomal fragile sites. Therefore, formation of R-loops contributes to the genomic instability observed upon replication stress.

1. INTRODUCTION

1.1 Genome instability: A driver of tumor progression

Genome instability has been identified as a hallmark of cancer [1]. Maintenance of genomic integrity in mammalian cells entails four major mechanisms: (i) high-fidelity of DNA replication in S-phase; (ii) accurate distribution of chromosomes between two daughter cells during mitosis; (iii) error-free repair of sporadic DNA damage throughout the cell cycle; and (iv) cell cycle progression and checkpoint control. In hereditary cancers, the cause of genome instability has been linked to mutations in DNA-repair genes. However, advancements in high-throughput sequencing technologies suggest that mutations in these caretaker genes do not account for the instability observed in sporadic cancers [2]. Most likely, this instability stems from **oncogene-induced replication stress**- impediments in the process of DNA replication that could lead to slowing, stalling or collapse of replication forks, mediated by activation of oncogenes.

1.2 Evidence for oncogene-induced replication stress

In line with the theory of oncogene-induced replication stress as an intrinsic driver of genomic instability, DNA damage could be detected during early stages of tumorigenesis in different cancers- bladder, breast colon and lung [3]. Increased replication stress may account for the observed damage in pre-cancerous lesions. Analyzing various stages of urinary bladder tumors revealed phosphorylation of DNA damage (DDR) proteins- Chk2 /ATM/γ-H2AX/ p53 in the initial phases of tumor development. Cultured cells over-expressing oncogenes like Cyclin E and Cdc25 have been shown to exhibit genomic instability and altered replication dynamics [4, 5]. According to the oncogene-induced DNA damage model of tumor development, replication stress activates DDR in pre-cancerous lesions, thereby activating checkpoints to promote apoptosis and senescence, creating a barrier against tumor progression. However, continuous DNA damage eventually generates a selective pressure for acquisition of mutations in tumor suppressors like p53. These mutations help breach this barrier and push cells towards cancer development [6](Figure 1).

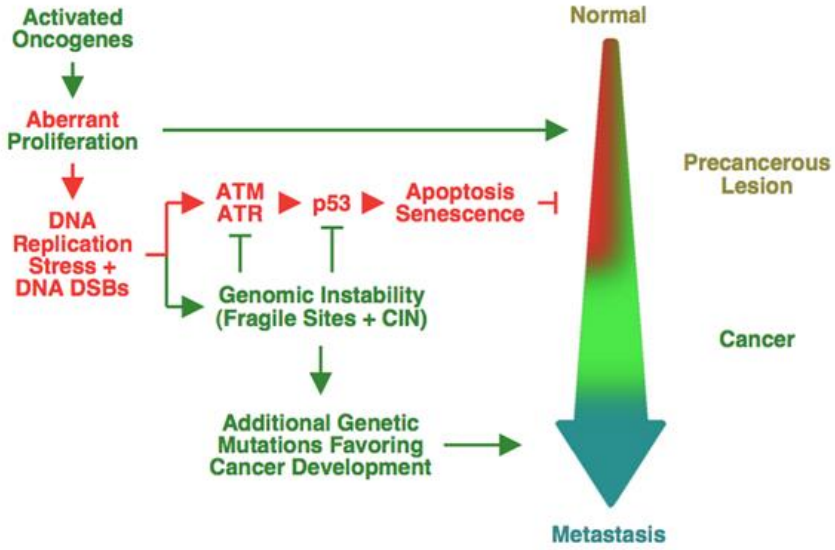


Figure 1. Model for oncogene-induced DNA damage and cancer development
Activated oncogenes induce replication stress generating DNA DSBs which activate *p53*, raising a barrier to tumor progression. Breach of this barrier by various mechanisms that impair the DNA damage response pathway allow cancers to develop [6].

1.3 Causes of oncogene-induced replication stress

How do oncogenes cause replication stress and trigger DNA damage? Although not very well established, molecular mechanisms responsible for the observed effects include- ***alteration of replication dynamics*** which may in turn lead to ***shortage of replication building blocks*** and ***collisions between transcription-replication complexes*** resulting in ***stable R-loop formation***.

1.3.1 Alteration of replication dynamics

Eukaryotes replicate their DNA from multiple origins distributed throughout their genome. The process of origin activation is temporally regulated to ascertain that the whole genome is duplicated in a given S phase and none origin is initiated for a second time in the same cell cycle. A two-step process ensures fidelity of origin activation: (i) origin licensing- loading of replicative helicase (Mcm2-7 complex) at putative origins; and (ii) origin firing- subsequent activation of the replicative helicase [7, 8]. Origin licensing occurs from late mitosis through G1 phase of the cell cycle,

while origin firing is restricted to the S phase, ensuring that once fired, an origin cannot be re-licensed, until it passes through the subsequent mitosis.

Origin under-usage

The licensing checkpoint ensures S phase entry only after all origins are licensed. However, upon oncogene over-expression in cancerous cells, the checkpoint may be compromised, leading to S phase entry with reduced origins. Although number of licensed origins far exceeds the ones fired, the back-up origins called dormant origins are necessary to rescue stalled forks upon replication stress [9]. Lack of licensed origins can cause DNA damage and genome instability in highly proliferative cancer cells as forks now have to travel greater distances to complete replication, increasing chances of fork stalling - lack of dormant origins does not allow rescue of the stalled forks, eventually leading to fork collapse and DSBs. The reduction in origin usage could also lead to cells entering mitosis with incompletely replicated DNA. This could cause defects in chromosome segregation and promote chromosome breakage [10]. Indeed evidence suggests that depletion of MCM increases the occurrence of anaphase bridges and micronuclei [11].

Over-expression of Cyclin E in human nasopharyngeal epidermoid carcinoma cell line, exhibited lower number of BrdU foci in the early S phase cells, indicating smaller number of DNA replication factories. This was accompanied by a reduction in chromatin binding of Mcm4 and Mcm7 in G1, indicating defective origin licensing [12].

Origin over-usage

Under normal conditions, most of the dormant origins are passively replicated during the S phase, and hence inactivated. However, over-expression of oncogenes may result in firing of otherwise dormant origins, as is seen upon Cyclin E, Myc and Ras overexpression, resulting in slower replication fork speed, increased fork stalling and reduced inter-origin distance, a sign of increased origin firing [13, 14].

Over-expression of oncogenes is known to cause pre-mature entry into S phase, shortening the G1 phase from 10-12 h to 2-4 h. According to a recent study, licensed origins within genes (intragenic) are inactivated by transcription during G1 [15]. Transcription through these regions strips off the replicative helicases, making them unable to fire during S phase. Oncogene over-expression and the resulting

short G1 do not allow transcription to complete before S phase entry, the novel origins are therefore fired within genomic domains normally devoid of replication initiation. Using genome-wide transcription and replication initiation studies, the authors could show that upon oncogene over-expression, a new class of intragenic origins fire, and these origins behave differently from the intergenic constitutive origins, as their replisomes are prone to fork collapse, and they are associated with the genomic rearrangements found in cancer [15] (Figure 2).

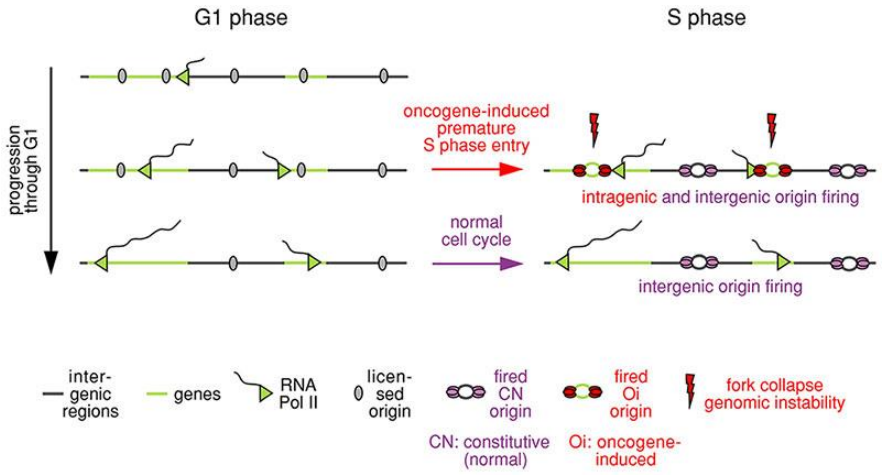


Figure 2. Mechanism of oncogene-induced replication stress
Shortening the length of the G1 phase of the cell cycle by activated oncogenes leads to firing of DNA replication within genes, collisions between the replication and transcription machineries, DNA breaks and genomic instability

1.3.2 Shortage of replication building blocks

Origin over-usage may exhaust substrates required for replication such as dNTPs. Replicative stress induced by some oncogenes can indeed be rescued by addition of exogenous nucleosides [13, 14]. Smooth DNA replication also requires synthesis of histones. Increase in the number of progressing forks could deplete cells of histones, thereby slowing fork progression, which is observed upon oncogene over-expression. Evidence shows that depleting proteins involved in histone biogenesis results in decreased fork speed, and chronic histone depletion eventually leads to DNA damage [16]. Other replisome proteins could also be a limiting factor. For example, RPA protects replication forks from nucleases by binding to ssDNA at ongoing forks and also to ssDNA generated at stalled replication forks. The

excessive DNA replication induced by oncogenes depletes RPA providing substrate for nucleases to generate DSBs [17].

An alternative hypothesis for depletion of replication factors could be the short G1 phase upon oncogene activation. Limited time for protein synthesis before S phase entry may yield insufficient material for successful replication, causing slow replication fork progression and increased fork stalling.

1.3.3 Transcription-replication collisions

Replication and transcription use the same DNA template; hence these processes routinely interfere with each other. In spite of a spatio-temporal segregation of these processes in eukaryotes, with DNA replication restricted to the S phase of the cell cycle while transcription of genes concentrated in G1, transcription-replication conflicts (TRCs) are still frequent events in certain genomic regions. The recent evidence that most of the genome is transcribed beyond gene boundaries, further enhances the probability of TRCs [18]. However, the directionality and functional state of the two machineries impact the consequence of TRCs. Studies indicate that head-on (HO) TRCs are more detrimental to fork progression than co-directional (CO) TRCs and result in genomic instability. This could be due to the fact that frontal clashes may cause RNAP complex to inactivate the replicative MCM helicase, blocking replication fork movement; or accumulation of positive supercoils between the two converging machineries can inhibit further DNA unwinding, hindering both transcription and replication [19]. CO conflicts, on the other hand, push RNAPII in the same direction as the replisome, positive supercoiling ahead of the replication fork neutralizes negative supercoiling generated behind RNAPII, allowing both processes to proceed. In fact, a preference for co-orientation of replication and transcription was observed in human cells, thus avoiding deleterious HO conflicts [20].

Perturbation of either replication or transcription programs can increase the possibility of head-on TRCs. As discussed earlier, activated oncogenes deregulate replication by changing origin usage patterns. In line with this, over-expression of Cyclin E was shown to increase origin firing and impair fork progression as well as increase DNA damage. These effects could be significantly reduced by inhibiting origin firing but also by inhibition of transcription, suggesting that TRCs contribute to

genome instability upon oncogene over-expression [13]. In yet another study, over-expression of the oncogene HRAS caused replication stress and TRCs by increasing global level of transcription via elevated expression of the general transcription factor TATA-box binding protein (TBP) [21]. Elevated TRCs upon oncogene over-expression could also be a consequence of the short G1 phase, as studies in drosophila, yeast and humans point to the ability of transcription to re-distribute replication origins in G1 outside transcription zones, thereby avoiding TRCs in the S phase [15, 22, 23] (Figure 2).

Another factor contributing to the deleterious effect of TRCs is further stabilization of the conflict via R-loop formation (Discussed in detail in section 1.4). A new study to address TRCs using an episomal system in human cells showed that HO-collisions promote R-loop formation and ATR activation, while CD-collisions seem to resolve existing R-loops. The CD bias of the genome thus keeps R-loops in check. However, replication slowdown or unscheduled origin firing (for example upon oncogene activation) increases HO-collisions as well as R-loop formation, resulting in genome instability [24].

1.3.3.1 Chromosomal fragile sites: Hot spots for TRCs

Chromosomal fragile sites are defined as regions of the genome which exhibit gaps or breaks on metaphase chromosomes under conditions of replication stress.

A well characterized category of fragile sites: Common fragile sites (CFSs) are expressed upon mild replication stress induced *in-vitro* by low doses of DNA polymerase inhibitor aphidicolin (APH). They occur in AT- rich, late replicating regions within large genes with a paucity of replication origins and have been associated with deletions in cancer [25]. Laszlo Tora's group demonstrated that these long genes that harbor CFSs require more than one cell cycle to complete transcription, so RNAPII transcribing in the late S phase could collide with perturbed replication forks, leading to the observed breakage [26]. Studies also propose CFS expression to be a consequence of active endonuclease mediated cleavage of unresolved replication intermediates in early mitosis rather than random mechanical chromosome breakage [27-29].

More recently, a new class of fragile sites- Early replicating fragile sites (ERFSs), have been identified within CpG rich early-replicating actively transcribed gene clusters, and are expressed upon conditions of replication arrest by hydroxyurea (HU) [30]. Although distinct, both types of fragile sites are protected by ATR kinase, and oncogene-induced replication stress triggers instability both ERFSs and CFSs, presumably as a consequence of TRCs [30, 31].

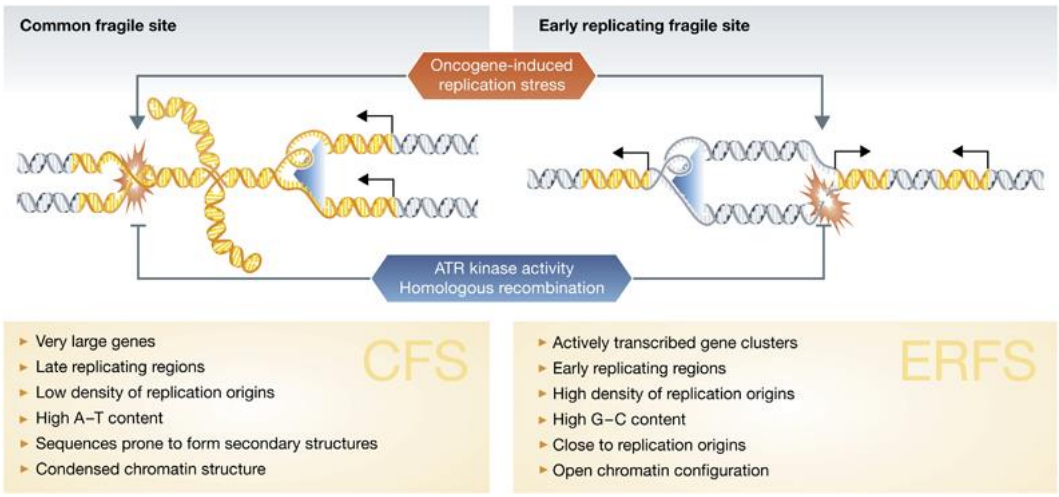


Figure 3. Comparison of CFSs with the newly identified ERFSs
Oncogene-induced replication stress causes replication fork stalling and collapse at both CFSs and ERFSs; ATR kinase and homologous recombination prevent collapse and mediate fork restart and repair

1.4 Co-transcriptional R-loops

R-loops are three- stranded nucleic acid structures comprising of an RNA:DNA hybrid and the displaced single-stranded non-template DNA strand. Short hybrids form physiologically during transcription within the RNA polymerase active site (8 bp) or during replication of the lagging strand (11 bp). However, longer and more stable hybrid tracts form co-transcriptionally when the mRNA exiting from a transcribing RNA polymerase re-anneals with the template DNA stand (Figure 4). Initially, R-loops were described in an *in-vitro* study that visualized their formation by electron microscopy [32]. Almost 20 years later, a study showed that R-loops could form *in vivo* during transcription in bacterial cells [33]. In the last decade, the field of R-loop

biology has expanded, with the advent of new methods to detect these structures *in vivo*.

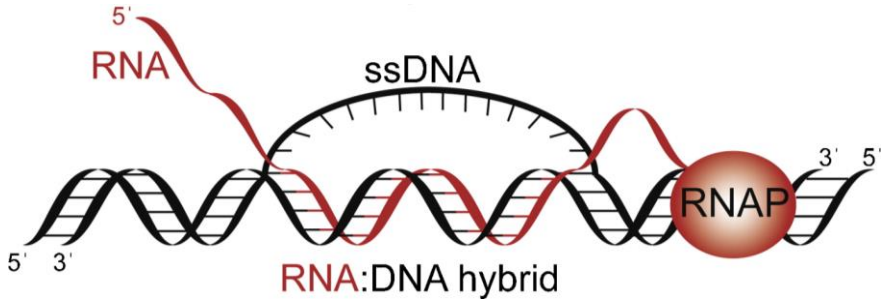


Figure 4. Structure of an R-loop
General structure of an R-loop: The nascent mRNA strand (red) synthesized by RNA polymerase (RNAP) hybridizes with the complementary DNA template strand. The non-template strand is exposed as single-stranded DNA (ssDNA) [34].

Different models were described to explain the formation of stable R-loops. The "extended hybrid model" claims that the 8 bp transient hybrid formed within the transcription complex can be extended to form an R-loop. However structural studies have ruled out this possibility due to the existence of separate channels for exiting of nascent mRNA and template DNA [35]. The most widely accepted model is the "thread-back model", in which the newly synthesized RNA exiting the RNA polymerase re-anneals with the template DNA, displacing the non-template strand, especially while transcribing through a C-rich template [36]. A third model describes the possibility of the nascent mRNA to re-anneal post-transcriptionally to a homologous sequence at a different locus, forming the R-loop in *trans* [37].

Formation of R-loops presents a competition between the nascent RNA and non-template DNA to hybridize with the template strand. Biochemically, a RNA:DNA hybrid is thermodynamically more stable than a DNA:DNA hybrid. R-loop stability is further enhanced when the non-template strand is G-rich. Strings of Gs on the non-template strand promote formation of G-quadraplexes, which further stabilize the ssDNA, facilitating RNA:DNA hybrid formation. Another factor favoring nascent mRNA hybridization is transcription-dependent DNA supercoiling. During transcription, positive supercoiling is generated ahead and negative supercoiling

behind the elongating RNA polymerase, thus allowing easy access to the mRNA when unresolved.

1.4.1 R-loops: Friends or foes?

Research so far indicates a paradoxical role of R-loops, on the one hand contributing to genomic instability and disease, while also playing an important part in various biological processes like transcription, DNA repair etc; discussed in detail in the following sections [38, 39] (Figure5).

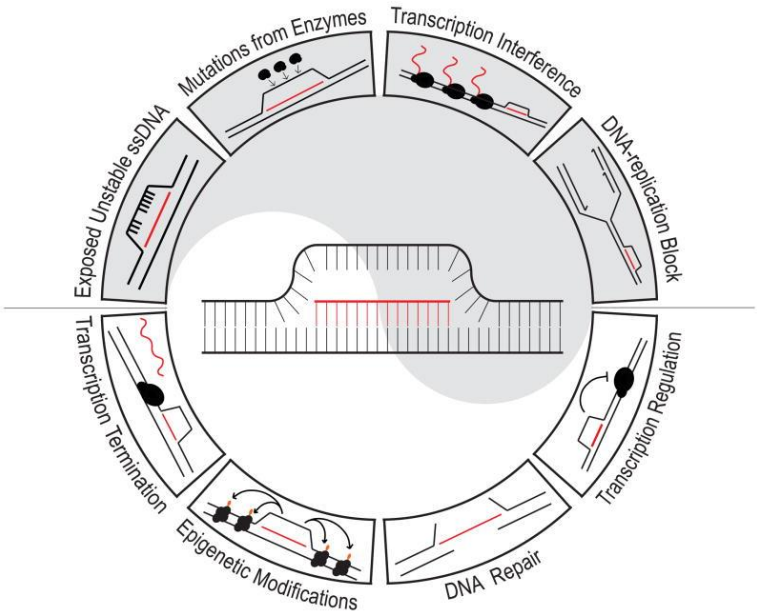


Figure 5. R-loops: A double-edged sword
R-loop formation can have a detrimental impact on genome stability- unstable ssDNA, transcription-replication conflicts or impaired replication progression but also contributes positively to regulate cellular processes like transcription and DNA repair [39]

1.4.1.1 Role in cellular processes: the "good" R-loops

A. Immunoglobulin class switch recombination

Antibody diversification requires the process of class-switch recombination (CSR), a mechanism that occurs in the immunoglobulin heavy (IgH) chain locus of mammalian B cells during immune response and mediates switching between the various immunoglobulin isotypes without affecting antigen specificity. The IgH locus is

known to have a high density of G-rich repetitive sequences. Studies have indicated that transcription through switch and constant regions precedes CSR and this is associated with co-transcriptional R-loop formation [40]. The action of activation-induced cytidine deaminase (AID) at the displaced G-rich ssDNA is a vital step for the generation of DSBs, which are then repaired by NHEJ, thereby mediating CSR [41]. This gave rise to a model wherein R-loops were proposed to enhance AID activity by providing it with abundant ssDNA substrate [42]. Furthermore, MCM helicase has been found to be enriched at IgH switch regions, where its helicase activity induces alterations in the local topology, in a manner dependent on transcription and R-loop formation. The topology favors DSB resolution by bringing DSBs in switch regions into proximity, promoting NHEJ. Reduction in R-loop formation decreased MCM complex loading, replication origin formation and CSR [43].

B. Transcription activation

CpG islands (CGIs) are found at the 5' end of genes, functioning as promoter elements for various genes including most of the housekeeping ones. Methylation of CGI promoters is very frequent and causes gene silencing, for example at genes expressed in a tissue-specific manner [44]. However, genome-wide studies have indicated that majority of the CGI promoters are unmethylated in normal tissues while aberrant methylation has been associated with various diseases like cancer [45, 46]. Thus, an important question has been to understand the mechanism underlying CGIs resistance to methylation, an otherwise abundant epigenetic modification. Various lines of evidence point to the role of transcription initiation at these loci in protecting against methylation. It was demonstrated that the presence of Pol II, active or stalled, predicts the epigenetic fate of promoter CGIs independently of transcription levels [47]. A co-relation between transcriptional output and protection against DNA methylation at promoter regions was also observed in the human methylome data. However, how transcription provides this protection is still not very well understood.

Genome-wide studies in human cells have marked the formation of R-loops at unmethylated CpG island promoters [48]. The formation of R-loops was shown to prevent CGI promoters from DNMT3B1-mediated DNA methylation. This was

attributed to R-loops serving as inappropriate substrates for DNMT3 activity or promoting recruitment of DNA demethylating complexes.

A recent study revealed that DNA methylase, DNMT1, favors binding and methylation of dsDNA over RNA:DNA hybrids. More than 1200 genes showed reduced methylation and active gene expression owing to R-loop formation. This was well demonstrated with the overexpression of RNase H1, that led to a significant increase in promoter methylation and thus lowered the gene expression pattern [49] (Figure 6).

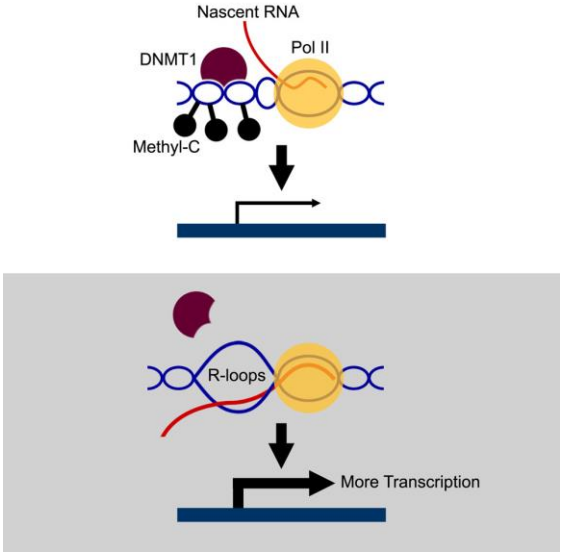


Figure 6. R-loops in transcriptional activation
Top: DNMT1 methylates cytosines at CGI promoters and prevents transcription.
Bottom: Formation of R-loops at promoters prevents DNMT1 binding to CGIs, and absence of methylation promotes transcription.

C. Transcription termination

Apart from a role at the 5' end of the genes, recent investigations have also detected the presence of R-loops at 3' end of ~2000 genes in human cells [50]. Here, R-loop formation at transcriptional pause regions was shown to halt RNA Pol II downstream of the poly(A) site, prior to termination. Over-expression of RNase H1 resulted in an increase in read-through RNA, indicating R-loop formation over pause elements to be an essential component of the termination process. Eventually these R-loops are removed, by the concerted action of a RNA /DNA helicase, Senataxin, and an exonuclease Xrn2, that release and degrade the nascent mRNA respectively, leading to successful transcription termination [51].

Formation of R-loops at termination sites has also been shown to induce antisense transcription at these regions resulting in dsRNA formation and recruitment of RNA-interference factors- DICER, AGO1, AGO2 and G9a histone lysine methyltransferase. This results in the establishment of an H3K9me2 repressive mark and recruitment of heterochromatin protein 1γ (HP1γ), thereby promoting localized chromatin condensation, impeding Pol II progression and efficiently terminating transcription [52].

D. Faithful chromosome segregation

ATR, a serine/threonine kinase, is a master regulator safeguarding the genome integrity. It is activated in response to replication stress and phosphorylates downstream targets to activate DNA damage checkpoints causing cell cycle arrest. The defined roles of this kinase have so far been restricted to the S phase. However, a recent study has shown ATR to be recruited to centromeres during mitosis in an R-loop dependent manner. Here, ATR inhibition in mitosis gave rise to lagging chromosomes, indicative of whole chromosome mis-segregation, distinct from persisting S-phase defects that represent as anaphase bridges. ATR inhibition also reduced levels of phosphorylated Aurora B, a protein that functions in mitosis to correct erroneous microtubule attachments at kinetochores. Phosphorylated ATR, CHK1, RPA as well as R-loops were also detected at centromeres. By expressing wild-type (resolve R-loops) and mutant forms (stabilize R-loops) of RNase H1, the authors demonstrated that R-loop formation presents RPA coated ss-DNA, activating ATR. ATR in turn activates CHK1, which promotes activation of Aurora B, thereby ensuring faithful chromosome segregation. This study provided a novel role of R-loops at centromeres in human cells, signaling accurate cell division [53].

E. DNA double-strand break repair

Cells are continuously challenged either by endogenous or exogenous sources of DNA damage. The DNA-repair machinery identifies and repairs the damage to maintain integrity of the genome. One of the most deleterious lesions encountered by the repair machinery is the DNA double-stranded break (DSB) which when inefficiently repaired, gives rise to serious chromosomal aberrations. To avoid

deleterious consequences arising from unresolved DSBs, cells employ two major pathways for DSB repair: non-homologous end joining (NHEJ) and homologous recombination (HR). While NHEJ, a more error-prone choice, is active throughout the cell cycle, HR is restricted to the S/G2 phases due to the requirement of homologous sister chromatids. This latter pathway is very well characterized, where the key players are known to be conserved among various species. HR requires the MRN complex (Mre11, Rad50, Nbs1) that first detects the breaks and activates ATM kinase, which phosphorylates various downstream substrates to initiate the DNA damage response (DDR). MRN together with CTIP mediate the initial resection of the broken DNA end generating 3'-overhangs. This is followed by the long-range resection of the 5' end by 5'-3' exonuclease EXO1 or by DNA2 nuclease that acts in conjunction with BLM or WRN helicases [54-56]. RPA is then recruited to coat and protect the resulting long ssDNA overhangs. Eventually, the RPA complex is replaced by RAD51 recombinase with the help of BRCA2 [57]. The RAD51 nucleofilaments mediate homology search and strand invasion into the sister chromatid to repair the damaged DNA [58].

In human cells, GFP-tagged RNase H1 was shown to be recruited to DSBs induced upon laser micro-irradiation in a manner dependent on transcription, indicating the formation of R-loops during DSB repair [59]. A later study in yeast demonstrated transient formation of RNA:DNA hybrids as an essential additional step in the classical HR-mediated repair pathway [60]. They revealed that RNA Pol II is recruited to the ssDNA generated upon resection around the DSB. Hence, the ssDNA generated upon resection during HR could potentially serve as a template for efficient transcription initiation. Although a significant increase in Pol II levels were detected around the break sites, this was not concomitant with a corresponding increase in transcript level, suggesting that the transcripts instantly hybridize with the template strand and remain trapped in RNA-DNA hybrid structures. (Figure 7)

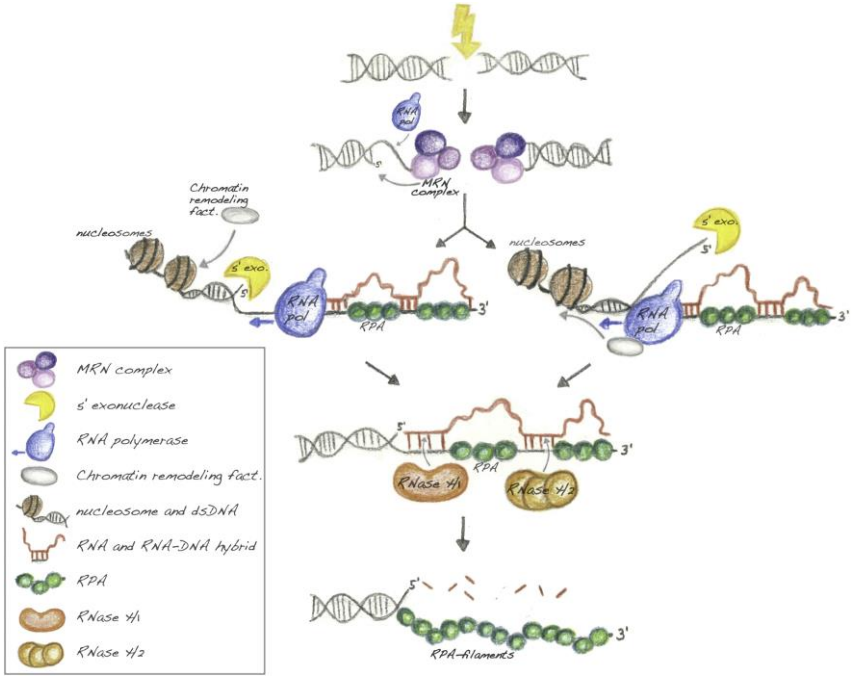


Figure 7. Model for transient RNA:DNA hybrids in DSB repair
RNA Pol II is recruited to the 3′ ssDNA overhangs and jump-starts transcription. The nascent RNA transcripts are prone to re-hybridize with the ssDNA template strand and form RNA-DNA hybrids, which directly compete with the ssDNA-binding RPA complex. Subsequent, long-range resection of the 5′ strand is performed by the 5′-3′ exonuclease Exo1.

Reducing the life time of the hybrids by RNase H1 over-expression results in the loss of genetic information around break sites as the transient formation of these hybrids seems to have a protective role in preventing unwanted intra-chromosomal recombination between repeat regions during HR-mediated repair. Apart from this, RNA DNA hybrids could also aid in choosing HR mediated repair over the error prone NHEJ. RNA:DNA hybrid formation at the break sites could prevent recognition of broken DNA ends by NHEJ proteins. In line with this view, a genome-wide analysis of DSB repair in human cells revealed HR as the preferred choice of DSB repair at transcriptionally active regions while NHEJ was favored at the transcriptionally inactive loci [61].

UV-induced photo lesions inhibit transcription by impeding polymerase progression. This causes displacement of co-transcriptional spliceosomes, which makes the unprocessed mRNA available to re-anneal with the template DNA. Formation of R-loops at these sites then activates ATM, which triggers a global displacement of late spliceosomes, affecting gene expression and pre-mRNA splicing genome-wide as part of the DNA damage response. Though more studies

into the exact mechanism are pending, this study depicts a non-canonical ATM signaling pathway, dependent on UV- induced R-loop formation [62].

1.4.1.2 Breaking bad: R-loops genome instability and cancer

R-loop formation, if not controlled, has been linked to different forms of genome instability- mutations, recombination, chromosome rearrangements as well as chromosome loss.R-loop formation renders the non-template DNA in single-stranded form. ssDNA is unstable and highly susceptible to DNA damaging agents. The enzyme AID, a cytidine deaminase, converts the dC residues to dU in the R-loops, which are then excised by the BER enzyme - uracil DNA glycosylase, creating an abasic site leading to ssDNA breaks [63]. Additionally, the endonucleases XPG and XPF have been shown to process R-loops into DSBs, demonstrating a role of TC-NER in promoting genomic instability [64]. In both yeast and human cells, R-loop accumulation has been associated with a concomitant increase in histone H3 phosphorylation (H3S10), a marker of chromatin condensation. It is intriguing to speculate that such R-loop mediated local chromatin compactions can induce genomic instability by serving as a barrier for replication fork progression [65]. Interestingly, a recent study identified histone mutants unable to acquire this phosphorylation incapable of inducing genomic instability upon R-loop formation [66]. Replication forks are further impaired when an R-loop displaces a G-rich ssDNA that forms G-quadraplexes. This stabilizes the existing RNA:DNA hybrid thus contributing to replication inhibition and instability [67].

Estrogen (E2) is a hormone that promotes the development of mammary tissue by stimulating transcription and driving breast cell proliferation. Elevated E2 levels are often associated with high incidences of breast cancer and are known to cause chromosome instabilities. A study revealed the stimulation of E2 levels in ER-positive breast cancer cells, to result in a rapid accumulation of co-transcriptional R-loops that could be detected at E2 responsive genes and also gave rise to DSBs in a manner dependent on replication. These loci were also significantly associated with breast cancer-specific chromosomal rearrangements, providing a link between R-loop formation and genomic instability observed in ER-responsive breast cancer cells [68]. BRCA1 and BRCA2, the two most frequently mutated tumor suppressor

genes in breast and ovarian cancers, have been shown to play a role in the prevention of R-loop formation and associated genomic instability [69, 70]. In yet another study, an oncogenic translocation identified in eosinophilic leukemia was shown to inactivate the polyadenylation factor FIP1L1, a protein associated with R-loop suppression [71].In yeast, Rad51 was identified to promote R-loop formation and genome instability, and RAD51 is also over-expressed in a number of cancers [37]. These observations raise the possibility that R-loops may provide proliferative advantages to tumor cells by promoting genome instability.

1.4.2 Limiting R-loop formation

Cells have to keep their R-loop levels in check in order to prevent the deleterious consequences of excessive R-loop formation. Many proteins from various pathways have been implicated in this process that either prevent R-loop formation or survey the genome for R-loops and resolve them when necessary. Some of these mechanisms are discussed in the following sections.

1.4.2.1 Factors preventing R-loops

A. RNA processing factors

Factors that prevent nascent mRNA to re-hybridize with template DNA would successfully prevent R-loop formation. Proteins involved in mRNA processing wrap nascent mRNA into ribonucleoprotein complexes, rendering them unavailable to form RNA:DNA hybrids. Evidence suggests that absence of co-transcriptional association of these proteins enhances R-loop formation and genome instability. The first studies came from yeast where mutants of hpr1 subunit of THO/TREX, a conserved protein complex involved in mRNA export, accumulated R-loops [72]. Additionally, they exhibited hyper-recombination phenotype and transcriptional arrest, which could be rescued by over-expression of RNase H1. The same group also extended their study to human cells and revealed that mutations in the THOC1 subunit of the human THO/TREX affects transcription elongation and mRNA export. They also observed R-loop-dependent genome instability and altered replication dynamics [73]. A more recent report identifies a physical interaction between THO and histone deacetlyase complex Sin3A. Depletion of THOC1/Sin3A causes hyper-

acetylation facilitating R-loop formation and replication fork stalling [74]. Along these lines, Aguilera's group further identified the nuclear basket Myosin-like protein 1 (Mlp1) as another factor involved in the prevention of R-loop formation in yeast [75]. Mlp1 is known to be required for a process termed as gene gating. According to the gene gating hypothesis, formation of an exportable mRNP is facilitated by transient localization of the transcribed RNA to the nuclear pore complex(NPC) [76]. R-loop formation in cells lacking Mlp1 could be suppressed by restoring physical proximity to the NPC, indicating these R-loops to be a consequence of defective gene gating. Although likely, such a mechanism is yet to be discovered in humans.

The most abundant class of RNA binding proteins in eukaryotes are the hetrogenous nuclear RNPs (hnRNPs) involved in mRNA processing, and arginine-rich proteins (SR family) involved in the splicing of mRNA precursors. Studies in DT-40 chicken cells and human Hela cells have shown the depletion of SR protein ASF/SF2 to result in chromosomal rearrangements and mutagenesis, which was rescued by RNase H1 over-expression [77]. In yeast, knock out of Npl3 hnRNP exhibited hypersensitivity to genotoxic agents and transcription-mediated genomic instability partly dependent of R-loop formation [78].

A plethora of new proteins involved in different steps of mRNA processing are being identified through various genetic/proteomic screens. One example involves an siRNA screen with γ-H2AX as a readout for DNA damage, which identified over 80 mRNP processing genes as suppressors of DNA damage in human cells. For a subset of these genes, over-expression of RNase H1 could rescue γ-H2AX accumulation [79]. Some noteworthy hits include helicase Aquarius, RNA splicing factors Skip, Cdc40 and snRNP proteins like Snrpa1.

B. Replication associated factors

***Topoisomerases*:** Various studies have shown the depletion of TOP1 to result in slower replication fork speeds and genomic instability, a phenomenon linked to the presence of R-loops [80]. An investigation in yeast showed the accumulation of R-loops at rRNA locus upon depletion of TOP1 [81].

Another topoisomerase, TOP3B, functions as a part of the TDRD-3 (Tudor domain-containing protein-3) by resolving negative supercoiling and R-loops during transcription, thereby preventing chromosomal translocations [82].

RPA: RPA, a ssDNA binding protein, is recruited to bind ssDNA at stalled replication forks or sites of DSBs, where it functions as a sensor to activate DNA damage response and DNA repair [83]. Interestingly, RPA is also detected during transcription, indicating the presence of ssDNA during the process, thus questioning its role in antagonizing R-loop formation [84]. To this end, experiments conducted by *Nguyen et al.* depicted an interaction between RPA and RNase H1, which when compromised, led to the accumulation of R-loops and associated genome instability [85]. In-vitro biochemical assays could further demonstrate the ability of RPA to stimulate the activity and binding of RNase H1 to RNA:DNA hybrids [85]. Taken together, these studies define a sensory role for RPA in R-loop recognition and resolution by RNase H1.

MCM-helicase: Besides its role in unwinding DNA during replication, biochemical studies indicated that MCM helicases could also unwind RNA:DNA hybrids *in vitro* while translocating on ssDNA, an activity that is conserved in bacteria, archaea and eukaryotes [86]. This brought attention to the role of MCMs in the resolution of R-loops. MCM also has other regulatory functions, most importantly in the activation of the DNA replication checkpoint. A mutation in MCM2 component of the helicase, mcm2DENQ, inhibited signaling upon replication stress, causing transcription-replication conflicts, thus inducing the formation of R-loops [87]. Origin usage patterns are affected if levels of MCM are perturbed, which may cause replication forks to approach transcription complexes in a different orientation. Indeed, depletion of MCM2 or MCM3 increases HO-collisions, resulting in increased R-loops and RNAPII-PCNA PLA foci, indicative of TRCs [24].

BRCA1/BRCA2: Mutations in BRCA1 and BRCA2 have been known to increase susceptibility to breast, ovarian, pancreatic and prostate cancers, reflecting the importance of their role in maintaining genome stability [88]. *Bhatia et al*. reported an interaction of BRCA2 with PCID2 (a subunit of the RNA export protein complex TREX-2), indicating a possible link to R-loop formation. They then concluded that

BRCA2 depletion increased the level of R-loops in cells, especially at actively transcribed genes. BRCA2-deficient pancreatic adenocarcinoma (CAPAN-1) cells exhibited pan-nuclear γ-H2AX staining that could be rescued by the over-expression of RNase H1. The authors proposed that TREX-2 may participate in the recruitment of BRCA2 to the proximity of transcribed genes, where it can bind the displaced ssDNA of the R-loop, allowing easy access for RNA-DNA hybrid resolution by RNase H1 or Senataxin [70].

While identifying novel binding partners of BRCA1 to help elucidate its role in cancer suppression, the Livingstone group showed an interaction between BRCA1 and Senataxin. BRCA1 depletion impaired SETX binding to transcriptional termination regions, suggesting a role for BRCA1 in serving as a scaffold for the recruitment of SETX. Depletion of SETX or BRCA1 increased R-loops and γ-H2AX accumulation at these regions [69]. ChIP data for BRCA1 show 764 distinct BRCA1 peaks overlapping significantly with transcription termination regions (TR) and 70% overlap with paused RNAPII indicating the involvement of BRAC1 in transcriptional termination. These sites are also enriched for R-loops as demonstrated by genome-wide DRIP analysis [50]. The screening of BRCA1 breast cancers revealed a significant enrichment for insertion-deletion mutations in the vicinity of the identified BRCA1 termination regions, further pointing to the role of this protein in regulating R-loop-mediated DNA damage.

Fanconi anemia proteins: The cancer predisposing syndrome- Fanconi anemia(FA), results from mutations in any of the 21 genes involved in the FA pathway. The most well characterized roles of FA proteins is in interstrand crosslink (ICL) repair as well as in stabilization of stalled replication forks. More recently, two major studies focused on their role in R-loop suppression. Depletion of FANCD2 and FANCA (mediates ubiquitination of FANCD2) increased the level of R-loops in cells and also displayed signs of genome instability. Interestingly, the authors also demonstrated a role for the translocase activity of FANCM in suppression of R-loops which was further confirmed by biochemical studies demonstrating that FANCM could unwind RNA:DNA hybrids in vitro [89]. Aguilera's group observed an increase in R-loop formation in FA patient cells defective in FANCA or FANCD2 and primary bone marrow cells from FANCD2 deficient mice. They also showed DNA breaks in

FA cells to be R-loop dependent [90]. Another study implicated that FANCD2 is required for faithful DNA replication through CFSs [91]. R-loop accumulation at CFSs in FANCD2-deficient cells was shown to drive genome instability. This was confirmed in a later study where FANCD2 binding sites were identified at CFSs by genome-wide ChIP-seq analysis upon mild replication stress [92]. Thus, a number of proteins associated with replication limit R-loop formation.

C. Chromatin modifiers

FACT (facilitates chromatin transcription) is a chromatin-remodeling complex known to swap nucleosomes to facilitate transcription and replication. It promotes RNAPII movement by interacting with histones, causing nucleosome disruption and reassembly, ahead and behind the transcribing polymerase, respectively [93]. Human and yeast FACT complexes interact with replication proteins and FACT deficient cells show impaired fork progression [94]. DRIP analysis in these cells indicated enrichment of R-loops and associated genome instability, suggesting a role for this complex in chromatin reorganization at regions of transcription-replication conflicts allowing the replisome to overcome transcription-mediated obstacles including R-loops [95].

H3K9 methylation (H3K9me2 or H3K9me3) is a mark of eukaryotic transcriptional silencing of heterochromatin and is enriched at telomeres, pericentromeric heterochromatin and repetitive elements (REs). Using *C.elegans* as a model system, a recent study demonstrated an increase in the expression of REs in mutants lacking H3K methylation. Increased transcription through these regions was also accompanied by increased R-loop formation and hypersensitivity to replication stress (Figure 8). This eventually results in copy number variations, insertion and deletions; especially at the transcribed REs [96]. These data illustrate that proteins that modulate chromatin states play a role in R-loop suppression.

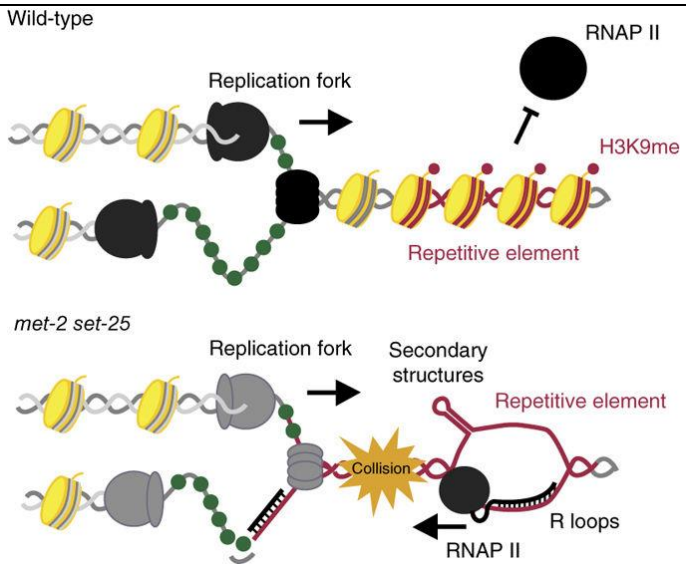


Figure 8: H3K9 methylation restricts transcription through repetitive elements The model illustrates how the loss of H3K9me could lead to the formation of secondary DNA structures and R-loops that engender replication stress specifically at heterochromatic repeats, to perturb genome integrity [96].

1.4.2.2 Re-solving the problem: Mechanisms to remove R-loops

Despite the existence of mechanisms to prevent R-loops, failure of these could lead to excessive R-loop formation. Also, the R-loops that play roles in cellular processes have to be removed in time. Cells employ nucleases or helicases to eliminate them, and this is discussed in the following section (Figure 9).

***Nucleases**:* RNase H enzymes are ribonucleases that degrade the RNA moiety of the RNA:DNA hybrid, allowing the two DNA strands to reassemble. Eukaryotes have two types of RNase H enzymes: RNase H1 and RNase H2. While both enzymes can degrade hybrids with more than 4 ribonucleotides, RNase H2 can also remove single ribonucleotides mis-incorporated during DNA replication [97]. In yeast strains lacking one or both RNase H enzymes, mitotic recombination events were mapped genome-wide. Deletion of RNase H1 does not cause deleterious effects, suggesting that RNase H2 can compensate for its function while double mutant displays high levels of recombination and chromosome instability [98]. However, in mice, RNase H1 knockout is embryonically lethal, owing to its role in mitochondrial replication and non-redundant functions of RNase H1 and RNase H2 in mammalian systems.

Knockdown of human RNase H1 causes defects in replication fork progression and leads to R-loop accumulation [99]. RNase H1 over-expression has been used to rescue genome instability associated with R-loop formation in a number of studies in bacteria, yeast and humans. Hypomorphic mutations in RNase H2 in humans can lead to the neurodegenerative disorder Aicardi-Goutiéres syndrome [100].

Helicases: Helicases can resolve R-loops by unwinding the RNA:DNA hybrid, allowing DNA strands to re-anneal.

Senataxin

The most well characterized helicase in R-loop resolution is Senataxin (Sen1), which was initially found as an RNAPII associated factor in yeast. A helicase dead mutant of Sen1 altered the distribution of RNAPII across the genome, indicating its role in regulating transcription [101]. Sen1 function was better understood when a study showed that it could resolve RNA:DNA hybrids, preventing transcription associated genomic instability [102]. Sen1 was also found to associate with replication forks, suggesting that it may help protect forks while they pass through highly transcribed genes [103]. Human Senataxin (SETX) was also shown to resolve hybrids, more specifically at transcription termination pause sites, where it is recruited by its interaction with BRCA1, and releases the RNA for degradation by exonuclease Xrn2 [51, 69]. The West group demonstrated that in response to replication stress, SETX co-localizes with DNA damage markers in a manner dependent on transcription and R-loop formation, suggesting that it localizes at the interphase of transcription and replication [104]. Recent publications also point to the role of SETX in resolving the RNA:DNA hybrids formed during DSB repair in transcriptionally active genes. Absence of functional SETX causes persistence of RNA:DNA hybrids around the DSB and leads to chromosomal deletions. Besides resolving the hybrid, SETX promotes RAD51 recruitment, preventing random rejoining of distant DNA ends, thereby limiting translocations [105, 106].

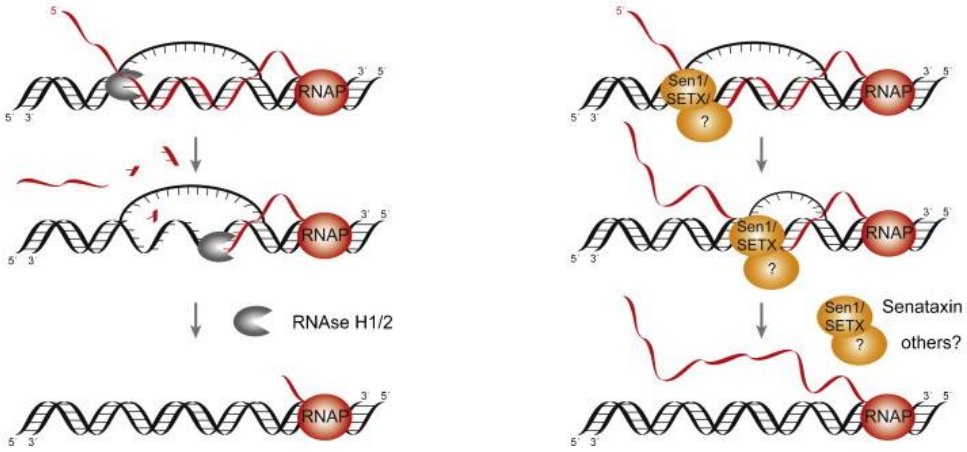


Figure 9. Mechanisms for resolution of R-loops
RNase H enzymes degrade the RNA moiety of the RNA:DNA hybrid while helicases like Senataxin specifically unwind the RNA:DNA hybrid and allow re-annealing of the non-template strand to restore the DNA double helix [34]

Pif1

Pif1 is a 5'-3', ATP dependent helicase, conserved from bacteria to humans. Yeast Pif1 has been shown to unwind RNA:DNA hybrids and maintain genome stability at G-quadruplexes [107, 108]. More recently, mechanistic insights revealed association of Pif1 to 3'ss-ds DNA junctions, where it induces a repetitive DNA looping coupled to its translocation activity. This DNA periodic patrolling mechanism of Pif1 is used to unwind RNA:DNA hybrids as well as G4-quadraplexes [109].

DEAD-box superfamily helicases

The nucleopore-associated mRNA export factor Ddx19, is a DEAD-box superfamily RNA helicase, recently shown to be recruited to the nucleus in response to replication stress, in a manner dependent on ATR/Chk1 activation, where it functions to resolve aberrant R-loops. The helicase can unwind hybrids *in vitro* and its depletion *in vivo* resulted in the accumulation of DNA damage as well as R-loops [110]. Another DEAD-box helicase, DDX21, was shown to be associated with PolI and PolII during transcription elongation and it can unwind co-transcriptional R-loops [111]. The Godbout group discovered an accumulation of DEAD box-1 protein (DDX1) at DSBs generated by ionizing radiation. By DRIP analysis, they could show that RNA:DNA hybrids persisted around DSBs upon DDX1 knockdown, indicating

that this helicase may be involved in removing the transient hybrids forming around the break [112] [113].

RecQ helicases

Among the RECQ family of helicases, Sgs1 and its human orthologue BLM were shown to unwind RNA:DNA hybrids *in vitro* [114]. Deletion of Sgs1 causes genome instability and R-loop formation, and regions of R-loop accumulation coincide with Sgs1 binding sites as well as the regions of chromosomal rearrangements. In human cells, BLM depletion causes R-loop-dependent genomic instability and this was also observed in BLM syndrome patient cells, further suggesting a possible involvement of these helicases in R-loop resolution [115].

1.4.3 Methods used to study R-loops

The S9.6 antibody, which detects RNA:DNA hybrids as small as 6 bp with an affinity of 0.6 nM, was initially developed to enhance DNA/RNA hybridization in DNA microarray studies [116]. However, in the last few years, it has been used widely as a tool to study R-loops *in vivo*. Staining with the antibody has been used to directly visualize R-loops in a number of studies in both mouse and human cells.

DNA:RNA immunoprecipitation (DRIP) with S9.6 antibody was used to pull-down R-loops from fragmented genomic DNA. R-loops at specific loci can be tested using q-PCR or genome-wide by next generation sequencing. Genome-wide profiling in yeast showed that R-loops form at rDNA, telomeres, transposons and a subset of ORFs with high GC content and high transcription. DRIP-Seq with cells expressing Sen1, RNase H1 and THO complex mutants revealed specific differences in R-loop formation. RNase H1 depletion led to R-loop accumulation at t-RNA and mitochondrial genes [117, 118]. Besides rDNA, genome-wide DRIP studies in human cells showed that R-loops are formed mostly at 5’ and 3’ ends (promoters and terminators) of genes in pluripotent embryonic cells [48, 50]. Another study in HEK293 and IMR-90 fibroblasts suggested that R-loops are distributed widely in intergenic regions, introns and repetitive DNA elements [119]. All these studies identified GC-skewness (asymmetry in nucleotide distribution) as a characteristic feature of R-loop formation. Intron-less genes harbor higher R-loop densities than

intron-containing genes of similar expression levels [120]. A high-resolution mapping of R-loops around promoters was achieved by combining bisulphite sequencing with DRIP- bisDRIP-seq, which uses bisulphite to convert cytosine to uracil residues within genomic regions containing ssDNA [121]. A computational algorithm to identify and map R-loop-forming sequences defined the need for a G cluster containing R-loop initiation zone (RIZ) and at its end, an R-loop elongation zone (REZ), with a high GC content. According to the database, oncogenes and tumor suppressors have the ability to form R-loops, linking R-loop formation to cancer [122].

Although, the S9.6 antibody is used widely, there have been inconsistencies in R-loop detection owing to its specificity. An RNase A pre-treatment was included to remove dsRNA in samples processed for S9.6 pull-down. A recent study showed that S9.6 could detect R-loops with a bias for certain sequences within the hybrid [123]. Understanding R-loop formation under different biological conditions requires the development of new tools and methods, and in this study, we sought to device a system to visualize and map R-loops under replication stress conditions in osteosarcoma (U-2OS) cells.

2. SCOPE OF THE STUDY

Formation of R-loops in cells has been associated with genome instability and cancer. R-loops can block the progression of replication forks and cause replication stress. However, if replication stress can promote R-loop formation is not well understood. The main aim of this study was to understand the formation of R-loops under conditions of replication stress.

We layed down the following major objectives:

1. Establish a tool to detect and isolate R-loops in U-2OS cells.
2. Using this tool, check for R-loop formation under different forms of replication stress. We focussed on mild replication stress mediated by low dose of aphidicolin and replication arrest induced with high dose of hydroxyurea.
3. Study the impact of R-loop formation upon replication stress on replication fork progression and its consequences in mitosis.
4. Genome-wide identification of R-loop forming loci under conditions of mild replication stress and replication arrest.

3. MATERIAL AND METHODS

3.1 Plasmid construction

The M27 variant of human RNase H1, lacking first 26 amino acids that contain mitochondrial targeting signal was used to construct WT and mutant RNase H1 versions Point mutation in the catalytic site of GFP-tagged hM27 RNase H1 (kindly provided by Dr. Robert Crouch), that generates catalytically dead enzyme [D210N (GAC→AAC)] [124], was introduced using Quick Change Site-Directed Mutagenesis kit (Agilent Technologies).

Following primers were used for mutagenesis:

RNH1-F(D210N): 5'-CTGGTTCTGTATACAACAGTATGTTTACGA

RNH1-R(D210N): 5'-TCGTAAACATACTGTTTGTATACAGAACCAG

The EcoRI-NotI fragment of plasmid pEGFP-N2 containing hM27 RNaseH1 (WT) or (D210N) was blunt ended by large Klenow fragment and subcloned into plasmid pAIO digested with EcoRV [125]. A DNA oligoduplex encoding for an shRNA for endogenous RNase H1 silencing was introduced between the BglII and HindIII sites of pAIO vector (shRNA targeted to region 730-751 RNaseH1).

Top strand RNase H1 sh730F:

5'GATCTACGATAAATGGTATAACTAACCTCGAGGTTAGTTATACCATTTATCGTTTTTTCTGCAGA

Bottom strand RNase H1 sh730R:

5'AGCTTCTGCAGAAAAAACGATAAATGGTATAACTAACCTCGAGGTTAGTTATACCATTTATCGTA

NotI-HindIII fragment of pAIO shRNase H1 containing shRNA was ligated with HindIII-NotI fragment of pAIO containing hM27 RNase H1-EGFP. In the resulting construct seven silent mutations were introduced into the RNase H1 cDNA between nucleotides 730-751 to render the resulting RNase H1 transcript resistant to the shRNA. Following primers were used for mutagenesis-

First round primers:

RH1_MTG730_1F:5'CAGACAGTATGTTTACCATCAACGGCATAACTAACTGGGTTCAAGG

RH1_MTG730_1R:5'CCTTGAACCCAGTTAGTTATGCCGTTGATGGTAAACATACTGTCTG

Second round primers:

RH1_MTG730_2F:5'GTTTACCATCAACGGCATCACGAATTGGGTTCAAGGTTGG

RH1_MTG730_2R:5'-CCAACCTTGAACCCAATTCGTGATGCCGTTGATGGTAAAC.

Underlined nucleotides were mutated.

3.2 Cell culture

U2OS T-REx derived cell lines were maintained in Dulbecco Modified Eagle's Medium (DMEM; Invitrogen) supplemented with Tet-free approved 10% fetal calf serum (FCS; Life Technologies), streptomycin/penicillin (100 U/mL), 1 µg/ml puromycin and 50 µg/ml hygromycin B (InvivoGen) at 37 °C in a humidified incubator containing 5% CO_2.

3.3 Inhibitors

The following compounds were used at the indicated final concentrations, unless stated otherwise: Doxycycline (1 ng/ml, Tocris), Aphidicolin (0.2 µM, Sigma-Aldrich), Hydroxyurea (10 mM, Sigma-Aldrich), Diospyrin D1 (5 µM; a kind gift from Dr. Banasri Hazra, Jadavpur University, Calcutta, India), Camptothecin (250 nM, Selleckchem), Olaparib (10 µM, Sigma-Aldrich), UCN-01 (300 nM, Sigma-Aldrich), Triptolide (1 µM, Sigma-Aldrich) Nocodazole (200 ng/ml, Sigma-Aldrich), R0-3306 (9 µM, Sigma-Aldrich), Cytochalasin B (2 µg/ml, Sigma-Aldrich)

3.4 Generation of stable cell lines

U2OS T-REx cells were transfected with 2 µg of pAIO-based constructs of GFP-tagged wild-type and catalytically-inactive mutant (D210N) of RNaseH1 and selected in the presence of 1 µg/ml puromycin (InvivoGen). Post-culturing for 10-14 days, healthy clones were isolated and positive clones were identified by testing for expression of the desired proteins by western blotting using anti-GFP antibody, before and after doxycycline induction.

3.5 Western blot analysis

Cells were trypsinized and washed with PBS before re-suspension in lysis buffer [50 mM Tris-HCl buffer (pH 7.5) ,120 mM NaCl, 20 mM NaF, 1 mM EDTA, 6 mM EGTA, 15 mM Na-Pyrophosphate and 0.5% NP-40] and sonicated for 5 min with a Diagenode sonicator. Cell lysate was clarified by centrifugation at 13000 rpm for 30 min at 4°C. Protein concentration of the soluble fraction was measured by Bradford's method. Samples were boiled in Laemmli sample buffer and separated by SDS-PAGE. Proteins were then transferred onto a Hybond-P PVDF membrane (GE Healthcare) in a semi-dry transfer apparatus at 56 mA for 90 min. Afterwards, the membrane was blocked with 5% milk in TBST (Tris-buffered saline supplemented with 0.01% Tween-20) for 30 min and then incubated with the respective primary antibodies (diluted in 5% milk/TBST) at 4°C overnight. The membrane was then washed three times in TBST and incubated with appropriate horseradish peroxidase-coupled (HRP) secondary antibody (anti-mouse 1:10000 and anti-rabbit 1:5000 dilution) for 1 h at RT. Then, the membrane was washed three times with TBS-T and bands were detected using ECL western blotting substrate (Pierce). The following primary antibodies were used for immunoblotting: GFP rabbit polyclonal (Abcam,1:2000), TFIIH rabbit polyclonal (Santa Cruz, 1:1000).

3.6 Immunofluorescence assays

Cells were grown on cover slips. After treatment with drugs, cells were washed with PBS (Phosphate-buffered saline) and pre-extracted using CSK buffer (25 mM HEPES, 50 mM NaCl, 1 mM EDTA, 3 mM $MgCl_2$,300 mM sucrose and 0.5% Triton X-100) on ice for 10min and then fixed with 4% formaldehyde to retain only chromatin-bound proteins. After several washes with PBS, cells were permeabilized with 0.5% Triton X-100/PBS for 5 min at RT. Cells were blocked in 5% BSA/PBS solution for 30 min and then incubated with the appropriate primary antibody (diluted in 5% BSA) at 4°C overnight. Cover slips were washed three times with PBS and incubated with the secondary antibody for 1 h at RT. After washing with PBS, cover slips were stained for 10 min at RT in the dark with 1 mg/ml DAPI (Sigma- Aldrich) diluted 1:2000 in PBS and mounted on glass sides using Fluoromount-G (Invitrogen). The mounted slides were left to dry overnight at 4°C. Slides were

analyzed with a fluorescent microscope at 63X magnification (Leica microscope, model DM6B, coupled to the DMC 2900 digital camera). The following primary antibodies were used for immunofluorescence staining: 53BP1 rabbit polyclonal (Santa Cruz; 1:200 dilution); Phospho-H2AX (Ser 139) mouse monoclonal (Millipore; 1:500 dilution). The secondary antibodies used were: Alexa Fluor 488 Goat Anti-Rabbit IgG (Thermo Fisher Scientific; 1:300) and Alexa Fluor 594 Goat Anti-Mouse IgG (Thermo Fisher Scientific; 1:300).

For measuring 53BP1 nuclear bodies in G1, S-phase cells were marked by pulse-labeling with EdU for 30 min; incorporated EdU was visualized by the Click-iT EdU reaction according to the manufacturer's protocol (Thermo Fisher Scientific).

3.7 Analysis of anaphase bridges and micronuclei

For anaphase bridge analysis, cells grown on coverslips were synchronized with 9 μM RO-3306 for 16 h, washed twice with 1X PBS for 5 min at RT and released into fresh DMEM medium for 1 h at 37°C. After release, cells were crosslinked with 4% formaldehyde for 10 min at RT, washed with PBS, followed by DAPI staining and mounting on glass slides. For quantification using a Leica DM6B fluorescent microscope, all anaphase cells were counted, and scored for the presence of DAPI-stained bulky bridges.

For micronuclei detection, cells were seeded on cover-slips and the medium was supplemented with 2 μg/ml cytochalasin B (micro-filament assembly inhibitor), for the last 16 h, to block cytokinesis in cells. These cells were then fixed with 4% formaldehyde for 10 min at RT and stained with DAPI. Images were acquired using a Leica DM6B fluorescent microscope at 63X magnification, and the percentage of bi-nucleated cells with micronuclei was determined.

3.8 Chromosome spread analysis

Asynchronously growing cells were supplemented with medium containing 200 ng/ml nocodazole, 5 h before harvesting by mitotic shake-off. Cells were collected in a 15 ml tube and centrifuged for 5 min at 4°C. The supernatant was removed and the cell pellet was resuspended in 1 mL of DMEM. To this, 8 mL of a 75 mM KCl solution was added and cells were incubated at 37°C for 15 min. 5 mL of freshly prepared

Carnoy's buffer (75% methanol, 25% glacial acetic acid) was then added to the cells in KCl solution for 15 min at RT. The fixing step using Carnoy's buffer was repeated thrice, re-suspending cells in 8 mL of Carnoy's buffer for each incubation. Cells were then spread drop-wise, from about 6 inches height, onto glass slides and dried overnight at RT. Next day, slides were stained with DAPI and mounted with coverslips using Fluoromount-G. Chromosome spreads were analyzed on a Leica DM6B fluorescent microscope at 100X magnification and visible gaps/breaks were quantified for each spread.

3.9 DNA fiber analysis

Cells were pulse-labeled with 30 µM of thymidine analogue chlorodeoxyuridine (CldU; Sigma-Aldrich) for 30 min, washed twice with PBS, followed by pulse labeling with 250 µM 5-iodo-2′-deoxyuridine (IdU; Sigma-Aldrich) for 30 min. Post-labelling, cells were washed twice with PBS, trypsinized and re-suspended in PBS to a concentration of 2.5×10^5 cells/mL. Labeled cells were diluted 1:1 with unlabeled cells, and 2.5 µl of cells were mixed with 7.5 µl of lysis buffer (200 mM Tris-HCl, pH 7.5, 50 mM EDTA, and 0.5% SDS) on a glass slide and incubated for 9 min. The slides were then tilted at 15–45°, and the spreads were air-dried and fixed at 4°C overnight in 3:1 methanol/acetic acid. DNA fibers were denatured with 2.5 M HCl for 1 h, washed with PBS, and blocked with 2% BSA in PBST (PBS supplemented with 0.1%Tween-20) for 30 min. Slides were incubated with primary antibodies against CldU (rat; Abcam) and IdU (mouse; Becton Dickson), respectively, for 2.5 h at RT. After five PBST (PBS supplemented with 0.2%Tween-20) washes, slides were incubated for 2 h in the dark at RT with the following secondary antibodies: anti–mouse Alexa Fluor 488 (Invitrogen) and anti–rat Cy3 (Immuno Research). The slides were again washed five times with PBST, air dried and mounted with 20 µl of Antifade gold (Invitrogen). Images were acquired with a fluorescence microscope (Leica microscope, model DM6B, coupled to the DMC 2900 digital camera) at 63X magnification. Replication tracts were measured using ImageJ software. Data were plotted using GraphPad Prism (Box and whiskers plot, whiskers drawn down to the 10^{th} percentile and up to the 90^{th}). Statistical significance was calculated using the Mann-Whitney test.

3.10 Chromatin Immunoprecipitation (ChIP)

Cells grown in 15 cm plates were pre-extracted using CSK buffer (25 mM HEPES, 50 mM NaCl, 1 mM EDTA, 3 mM $MgCl_2$,300 mM sucrose and 0.5% Triton X-100) and crosslinked with 1% (v/v) formaldehyde for 15 min at 30°C. Fixation was stopped by adding glycine to a final concentration of 125 mM followed by incubation for 10 min at 30°C. After washing twice with PBS, cells were collected with a cell scraper in 1 mL of cell lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% NP-40 and 1 X protease inhibitor cocktail) and incubated for 30 min on ice. Nuclei were collected by centrifugation at 5000 rpm for 10 min and suspended in nuclear lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS and 1 X protease inhibitor cocktail). Chromatin DNA was sheared to ~300 bp fragments by sonication using the Diagenode bioruptor on high intensity for 30 min. To quantify the chromatin, a 50µl aliqot of fragmented DNA was subjected to overnight reverse cross-linking at 65°C using 300mM NaCl. Next day, following RNase A treatment for 15 min at 37°C and deproteinisation using Proteinase K at 55°C for 1 h, DNA was purified and quantified using a nanodrop. 10 µg of chromatin was then incubated with magnetic beads conjugated with anti-GFP antibody (rabbit polyclonal, Abcam) overnight in IP buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.5% NP-40, 1% Triton X-100 and 1 X protease inhibitor cocktail) at 4°C. 5% of fragmented chromatin was saved as input. The following day, beads were sequentially washed three times with IP buffer, low salt buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA and 1 X protease inhibitor cocktail), high salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA and 1 X protease inhibitor cocktail), once with LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1% NP-40, 1% Deoxycholate, 1 mM EDTA and 1 X protease inhibitor cocktail) and once with TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA). The protein-chromatin complex was eluted with elution buffer (1% SDS, 300mM NaCl and 0.1M $NaHCO_3$) and reverse-crosslinked by incubation overnight at 65°C. After sequential RNase A and Proteinase K (0.2 µg/mL) treatment at 55°C for 1 h, DNA fragments were purified using QIAquick PCR purification kit (QIAGEN). DNA was then either used for qPCR analysis or for library preparation followed by ChIP-seq.

3.11 DNA: RNA immunoprecipitation (DRIP)

Genomic DNA was extracted from U2OS T-REx cells by chloroform extraction, precipitated with ethanol and re-suspended in TE (10 mM Tris-HCl pH 8.0 and 1 mM EDTA). DNA was digested with a cocktail of restriction enzymes (BsrgI, EcoRI, HindIII, SspI, XbaI) overnight at 37 °C. For RNase H treated samples, 8 µg of digested DNA was treated with RNase H (NEB) overnight at 37 °C. For RNA:DNA hybrid immunoprecipitation, 4 µg of DNA was incubated with 10 µg of S9.6 antibody (HB8730) [116] overnight in 1X binding buffer (10 mM NaPO4 pH 7.0, 140 mM NaCl, 0.05% Triton X-100) at 4 °C. 5% of DNA was removed as input. The following day, 20 µL Protein A/G sepharose magnetic beads (Pierce) were added for 2 h at 4 °C. Beads were washed thrice with binding buffer and incubated with elution buffer (50 mM Tris pH 8, 10 mM EDTA, 0.5% SDS, Proteinase K) for 45 min at 55 °C. Eluted DNA was purified by phenol:chloroform extraction, followed by ethanol precipitation and subjected to qPCR analysis.

3.12 q-PCR analysis

Purified DNA samples obtained by ChIP and DRIP methods were analyzed by qPCR performed on a Roche LightCycler 480 Instrument II using SYBR-Green master mix (Roche). Relative enrichment was calculated as the amount of precipitated DNA relative to the amount of DNA present in input chromatin (% of input), and normalized to the control sample. The primers used for gPCR analyses are listed below:

Gene	Sequences of primers (in 5'→ 3' direction)	
	Forward (Fwd)	Reverse (Rev)
RSP19	GAGGCAGAGGTTGCAGTGAGTC	CTGGTAGAGAACAAGCTCCCAT
RPL22	GTGAACGGAAAAGCTGGGAAC	CTGGGCACTGGGTGCACGCA
ACTG1-3	GGTGACACAGCATCACTAAGGG	GACAGCACCGTGTTGGCGTA
ACTG1-5	GGTGACACAGTGAGACCCTATCT	GGCGTTCTTTACATATTGTGGAT
APOE	CCGGTGAGAAGCGCAGTCGG	CCCAAGCCCGACCCCGAGTA
RPL13A	GCTTCCAGCACAGGACAGGTAT	CACCCACTACCCGAGTTCAAG
BTBD19	CCCCAAAGGGTGGTGACTT	TTCACATTACCCAGACCAGACTGT

DHFR-2	GTTCTATAGTCACTGCATCTTAGTC	TGCTAATTCTGGTTGTTCAGTAAG
DHFR-4	TTGTTTCAGGGACAGGGTCTT	CTGTGGTGGGAAGATGGCT
FRA3B	CACTTCCTAACAGGCCCAAA	CCTCCACTTCTCCTCCCTCT
FRA16D	TCCTGTGGAAGGGATATTTA	CCCCTCATATTCTGCTTCTA

3.13 ChIP-sequencing and analysis

Purified DNA samples obtained by ChIP were quantified with Qubit dsDNA HS assay (Invitrogen) and fragment size distribution was monitored by capillary electrophoresis (Agilent 2100 Bioanalyzer, High Sensitivity DNA Chips (Agilent), and sequenced according to Illumina protocols. Briefly, sequencing libraries were prepared using the NEBNext Ultra DNA Library Prep kit for Illumina (E7370S, NEB). Following adaptor ligation, DNA was PCR amplified and libraries were sequenced paired-end on a HiSeq 2500 instrument (Illumina).

Sequencing data generated was first subjected to quality checks using FastQC analysis. Trimmomatic and Cutadapt were used to trim the reads and remove low-quality and adaptor sequences [126]. Processed reads were mapped to GRCh38 standard reference assembly using GSNAP[127]. Peak calling was performed using MACS2 [128] and suspicious peaks were defined as the union of all peaks from the biological replicates and treatments, enriched over the input requiring at least twofold enrichment and $q < 0.05$. DESeq2 was used to obtain among the suspicious peaks the treatment specific peaks, defined by two-fold enrichment between treatment groups but similar coverage within the same treatment. Statistical significance was controlled by false discovery rate (FDR < 0.1) [129]. Bedtools was used to overlap peaks to regions in the genome and annotations were in line with EnsEMBL version 90 [130]. SortMeRNA was used to quantify fraction of rDNA reads [131]. Plots were produced with deepTools package or in R.

4. RESULTS

4.1 Catalytically-inactive RNase H1 can be used to detect R-loops in human cells

The monoclonal antibody (S9.6) that recognizes RNA:DNA hybrids has been widely used to study R-loops in cells[116]. We established another tool- a stable U2OS-T-REx cell line that upon induction with doxycycline (dox) expresses catalytically-inactive RNase H1 (D210N) as a fusion with green fluorescence protein (GFP) (Figure 10A). This RNase H1 mutant can efficiently bind to R-loops but is unable to degrade them ([124]. The use of a GFP-tag enables the detection of R-loops by fluorescence microscopy and allows their isolation by immunoprecipitation using anti-GFP antibody. We also generated an U2OS-T-REx cell line for inducible expression of GFP-tagged wild-type RNase H1 (Figure 10B).

To demonstrate that this cellular system could be used for detection of R-loops, cells were exposed to diospyrin D1, an inhibitor of the spliceosome assembly known to induce R-loop formation [80]. To detect binding of RNH1(D210N)-GFP to chromatin, cells were pre-extracted to remove unbound proteins and analyzed by fluorescence microscopy. We observed nuclear retention of RNH1(D210N)-GFP in majority of diospyrin-treated cells, indicative of R-loop formation (Figure 10C, D). Mock-treated cells displayed nucleolar a RNH1(D210N)-GFP signal, indicative of R-loop formation in rDNA regions (Figure 10C).

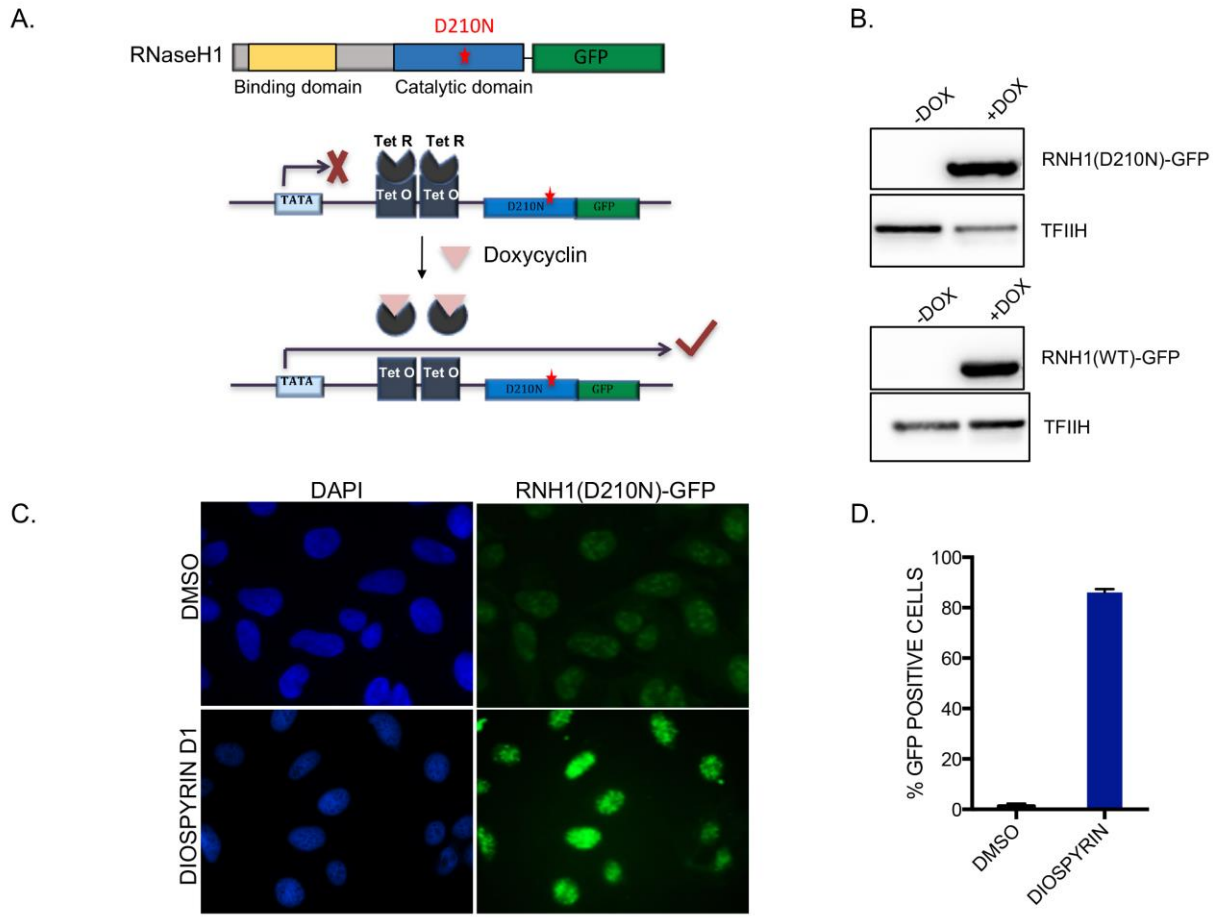


Figure 10: Establishment of a tool to detect R-loops

A. A scheme of RNase H1 fused to green fluorescence protein (GFP): position of the D210N substitution in the catalytic domain of RNase H1 is indicated. This mutant of RNase H1 can bind to RNA:DNA hybrids, but fails to degrade these structures (*top panel*).
Schematic of the inducible system for expression of RNH1(D210N)-GFP. RNH1(D210N)-GFP was cloned to the plasmid pAll-In-One (pAIO) under the control of a CMV promoter containing two copies of the tet operator (TetO2) sequence. This plasmid was transfected into U2OS T-REx cells to generate a stable cell line. Tet repressor, which is constitutively expressed in U2OS T-REx cells, binds to tet operator sequence and blocks expression of RNH1(D210N)-GFP. Addition of doxycyclin (dox), which binds with high affinity to the tet repressor and causes a conformational change that releases it from the tet operator, expression of RNH1(D210N)-GFP is induced (*bottom panel*).
B. Western blot analysis of extracts of U2OS T-REx cells expressing RNH1(D210N)-GFP and RNH1(WT)-GFP upon induction with 1ng/ml dox for 24 h.
C. Representative images of cells expressing RNH1(D210N)-GFP induced by 1 ng/ml dox for 24 h, and treated with either DMSO (mock) or 5 µM diospyrin D1 for the last 4 h. After pre-extraction, cells were fixed with formaldehyde and nuclei were stained with DAPI.
D. Quantification of images represented in C, depicted as percentage of GFP-positive cells. Data are represented as mean of three independent experiments; error bars show SEM.

4.2 R-loops are formed in response to replication stress

Having established the tool to detect R-loops in cells, we wanted to test for R-loop formation under conditions of replication stress (RS). Cells expressing mutant RNase H1 were exposed to (i) low dose of aphidicolin (APH- a DNA polymerase inhibitor); and (ii) high dose of hydroxyurea (HU- causes dNTP depletion by inhibiting ribonucleotide reductase). We observed that a significant fraction of cells formed R-loops under these conditions. These cells were also positive for the histone variant γ-H2AX, indicative of RS (Figure 11A, B). We then performed a time course analysis to determine the duration of these treatments that induces maximum R-loop formation. Prolonged treatment with APH (16 h) and rather short exposure to HU (6 h) yielded the highest level of R-loops (Figure 11C, D). Other forms of RS were also tested for their ability to promote R-loop formation. Camptothecin (CPT)- a topoisomerase I inhibitor and UCN-01, inhibitor of CHK1, generated around 40% cells positive for R-loops and also γ-H2AX (Figure 11E, F).

Next, we wanted to check the effect of transcription inhibition on RS induced R-loop formation. We treated cells with Triptolide (TRP), an inhibitor of transcription initiation, prior to and during HU treatment, and observed a significant decrease in R-loop formation, suggesting that TRCs upon RS may result in the formation of R-loops (Figure 11G).

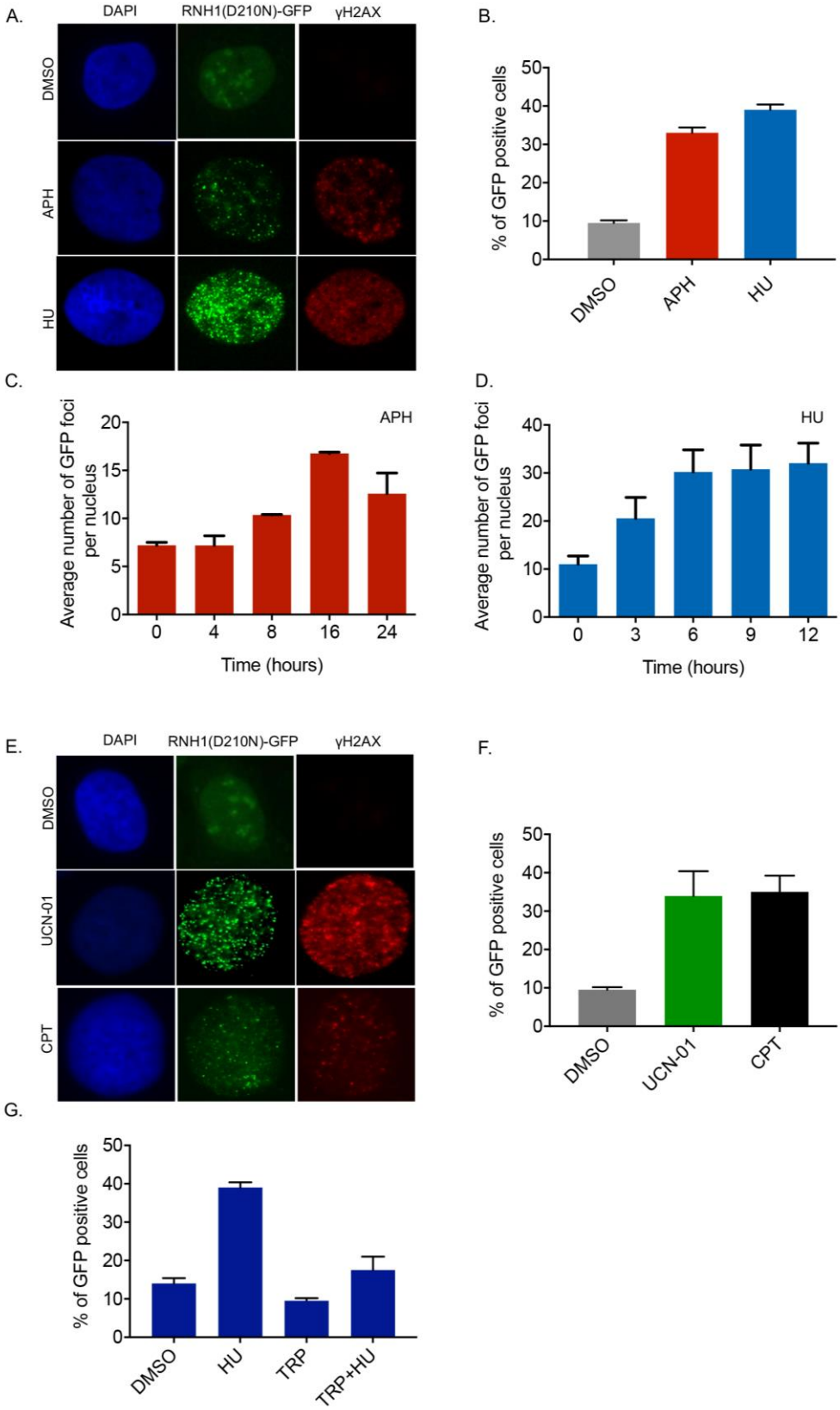


Figure 11: R-loop formation upon replication stress

A. Representative immunofluorescence images of U2OS T-REx [RNH1(D210N)-GFP] cells induced with 1 ng/ml dox for 24 h and treated with DMSO (mock) or APH (0.2 μM) for the last 16 h; or HU (10 mM) for the last 6 h. Cells were pre-extracted to detect chromatin bound protein and immunostained for γ-H2AX. Nuclei were stained with DAPI.

B. Quantification of GFP-positive cells in A. Data are represented as mean of three independent experiments; error bars show SEM.
C. Average number of RNH1(D210N)-GFP foci per nucleus of cells induced with 1 ng/ml dox for 24 h, and co-treated with APH (0.2 μM) for the indicated time points.
D. Average number of RNH1(D210N)-GFP foci per nucleus of cells induced with 1 ng/ml dox for 24 h, and co-treated with HU (10 mM) for the indicated time points.
E. Representative immunofluorescence images of U2OS T-REx [RNH1(D210N)-GFP] cells induced with 1 ng/ml dox for 24 h and treated with UCN-01 (300 nM) for the last 6 h; or CPT (250 nM) for the last 16 h. Cells were pre-extracted to detect chromatin bound protein and immunostained for γ-H2AX. Nuclei were stained with DAPI.
F. Quantification of GFP-positive cells in E. Data are represented as mean of three independent experiments; error bars show SEM.
G. Quantification of GFP positive cells: U2OS T-REx [RNH1(D210N)-GFP] cells induced with 1 ng/ml dox for 24 h and treated with DMSO (mock); TRP (1 μM) 6 h; or HU (10 mM) for 5 h; or TRP (1 μM) for 1 h combined with HU (10 mM) for the last 5 h. Cells were pre-extracted to detect chromatin bound protein and nuclei were stained with DAPI.

4.3 R-loops pose a roadblock to replication fork progression under conditions of replication stress

Replication fork progression is impaired when cells are exposed to DNA synthesis inhibitors, either due to reduction in polymerase speed or presence of hurdles that obstruct polymerase movement [132]. To visualize individual forks, DNA fiber analysis was performed wherein cells were sequentially pulse-labeled with halogenated thymidine analogues CldU and IdU for 30 minutes each, and either a low dose of APH or high dose of HU were added during the second labeling. As expected, we observed significant shortening of the IdU tracts under these conditions, indicating fork slowing (Figure 12A). To check if R-loops contribute to this fork slowing phenotype, we over-expressed WT RNase H1 24 hours prior to DNA labeling. We could observe a partial rescue of replication fork progression upon RNase H1 over-expression both in APH- and HU-treated cells, indicating that R-loops retard fork movement under RS conditions. However, the observed IdU tract shortening was primarily the consequence of DNA polymerase inhibition by APH or dNTP depletion upon HU (Figure 12B).

To distinguish between the overall DNA replication slowing due to inhibition of DNA polymerase and fork stalling by a barrier such as an R-loop, we measured the progression of sister replication forks. In non-treated cells, sister forks progress at a similar rate from a given origin generating a symmetrical pattern in DNA fiber assay (ratio of the lengths of two IdU tracts~1), whereas on encountering an obstacle on

any one side, an IdU tract asymmetry can be observed (ratio of the shorter to the longer IdU tract <1). We could detect a significant increase in fork asymmetry when cells were exposed to APH or HU (Figure 12C). To test whether this replication tract asymmetry is a consequence of R-loop formation, we measured the progression of sister forks upon resolving R-loops by over-expression of RNase H1. We observed that both APH- and HU-induced sister fork asymmetry was completely suppressed upon RNase H1 overexpression, indicating that R-loops pose a roadblock to fork progression under RS conditions (Figure 12D).

Next, we wanted to test if these R-loops form as a consequence of collisions between transcription and replication machineries. We therefore sought to test whether inhibition of transcription affects replication fork progression under RS. To this end, cells were pre-treated with transcription inhibitor triptolide (TRP), two hours prior to pulse-labelling with halogenated nucleotides and second labeling was combined with either APH or HU treatment as in previous experiments (Figure 12E). Fork slowing upon RS was significantly rescued when transcription was inhibited (Figure 12F). In addition, sister fork asymmetry observed after APH and HU treatments was fully suppressed when transcription was inhibited prior to these treatments (Figure 12G). These data suggest that the block in fork progression is a consequence of TRCs, and if transcription is inhibited, fork stalling in response to RS can be rescued, presumably because R-loop formation is suppressed.

We also expressed the catalytically inactive mutant of RNase H1, which would stabilize rather than resolve R-loops, and performed the same DNA fiber assays. Over-expression of mutant RNase H1 further slowed replication fork progression upon APH treatment. Additionally, unlike WT RNase H1, the fork asymmetry pattern in APH-treated cells could not be rescued by over-producing the mutant enzyme. These results confirm that formation and persistence of R-loops upon RS impede replication fork progression (Figure 12H, I).

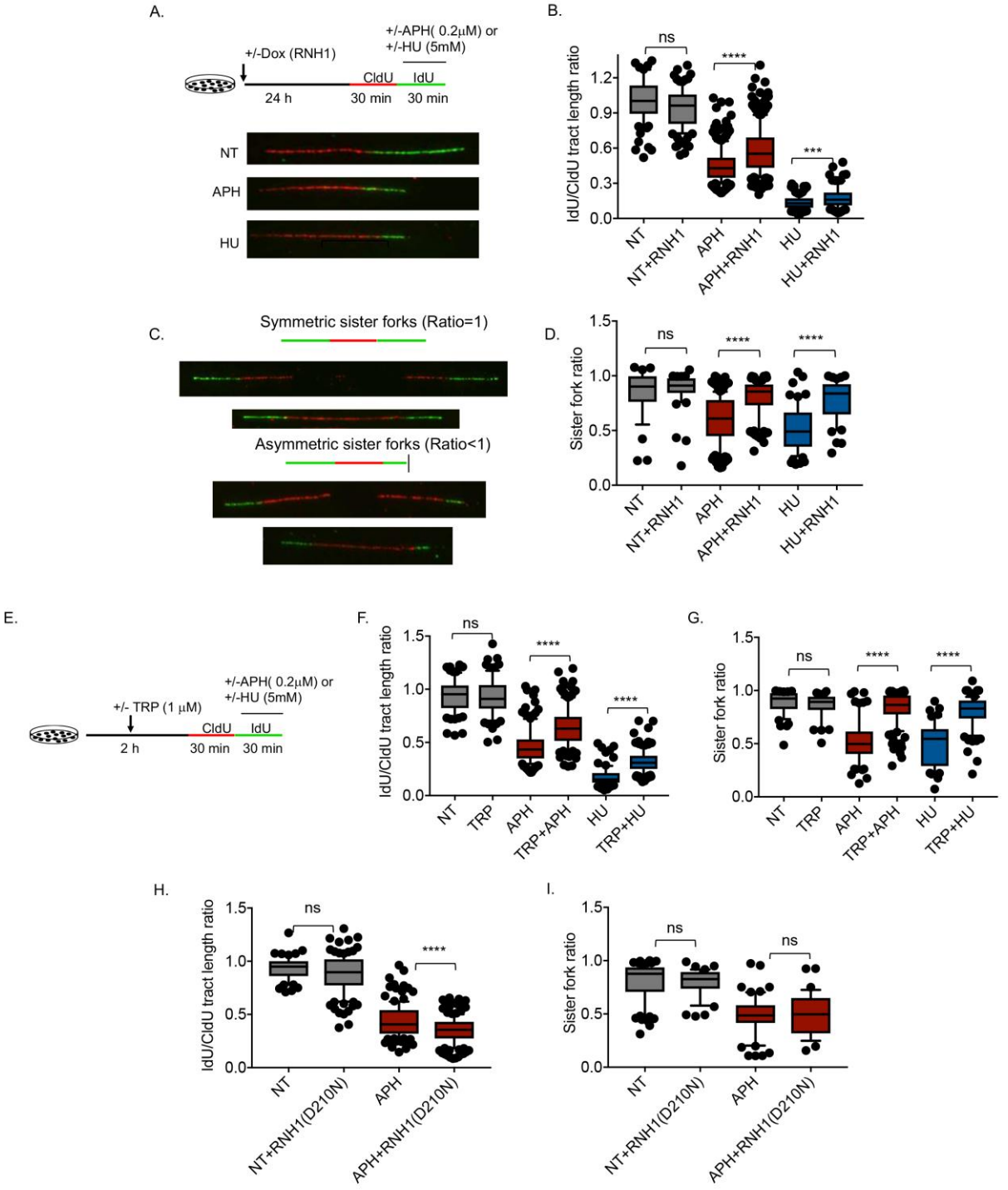


Figure 12: R-loops impair replication fork progression upon replication stress

A. Top panel: Labeling scheme for DNA fiber analysis: U2OS T-REx [RNH1-GFP] cells were labelled sequentially with CldU or IdU for 30 min each. Where indicated, APH (0.2 μM) or HU (5 mM) were added during the IdU pulse. For non-treated (NT) control, DMSO was added. Prior to the labelling, expression of WT(RNH1)-GFP was induced (RNH1) with 1 ng/ml Doxycycline (dox) for 24 h.

Bottom panel: Representative images of DNA replication tracts detected on DNA fibers upon the indicated conditions.
B. Ratio of the IdU to CldU tract lengths was plotted for indicated treatments following the experimental workflow described in (A). At least 200 replication tracts were scored per sample. Horizontal lines represent the median value, and boxes and whiskers show 10–90th percentiles. Statistical analysis according Mann–Whitney test: ns, not significant; *** P≤ 0.001; ****, P ≤ 0.0001.
C. Scheme and representative images of symmetric forks (Ratio of IdU tract lengths =1) and asymmetric forks (Ratio of IdU tract lengths <1)
D. Ratio of lengths of IdU tracts of sister forks were plotted for indicated treatments following the scheme described in (A). At least 150 sister forks were scored per sample. Horizontal lines represent the median value, and boxes and whiskers show 10–90th percentiles. Statistical analysis according Mann–Whitney test: ns, not significant; ****, P ≤ 0.0001.
E. Labeling scheme for DNA fiber analysis upon transcription inhibition: U2OS T-REx [RNH1-GFP] cells were labelled sequentially with CldU or IdU for 30 min each. Where indicated, APH (0.2 μM) or HU (5 mM) were added during the IdU pulse. For non-treated (NT) control, DMSO was added. Prior to the labelling, cells were treated with TRP (1 μM) for 2 h.
F. Ratio of the IdU to CldU tract lengths was plotted for indicated treatments following the experimental workflow described in (A). At least 200 replication tracts were scored per sample. Horizontal lines represent the median value, and boxes and whiskers show 10–90th percentiles. Statistical analysis according Mann–Whitney test: ns, not significant; *** P≤ 0.001; ****, P ≤ 0.0001.
G. Ratio of lengths of IdU tracts of sister forks were plotted for indicated treatments following the scheme described in (A). At least 100 sister forks were scored per sample. Horizontal lines represent the median value, and boxes and whiskers show 10–90th percentiles. Statistical analysis according Mann–Whitney test: ns, not significant; ****, P ≤ 0.0001.
H. DNA fiber analysis for cells expressing RNH1(D210N)-GFP upon treatment with 1ng/ml dox for 24hrs (D210N-RNH1). DNA fiber analysis was performed as in (A). Ratio of the IdU to CldU tract lengths was plotted. Horizontal lines represent the median value, and boxes and whiskers show 10–90th percentiles. Statistical analysis according Mann-Whitney test: ns, not significant; ****, P ≤ 0.0001.
I. Sister fork ratio upon expression of RNH1(D210N)-GFP. DNA fiber analysis was performed as in (C). At least 100 sister forks were scored per sample. Horizontal lines represent the median value, and boxes and whiskers show 10-90th percentiles. Statistical analysis according Mann–Whitney test: ns, not significant.

4.4 R-loops contribute to chromosome segregation impairment and genome instability in cells exposed to replication stress

We detected R-loop formation in cells upon mild RS (low dose of APH) and in addition, we showed that impairment of fork progression upon APH was dependent on R-loop formation. To further explore the effect of R-loop formation in cells when they enter mitosis, we scored for chromatid breaks on prometaphase chromosome spreads of cells exposed to APH. Low dose of APH stimulates chromosome fragility,

which is seen as appearance of approximately two breaks/gaps per chromosome spread. Resolution of R-loops by over-production of WT RNase H1 significantly reduced the formation of these breaks/gaps while stabilization of R-loops by over-expression of the mutant form of RNase H1 (D210N) further increased the fragility (Figure 13A).

We then asked if persistent R-loops could cause chromosome segregation defects upon RS. The increase in DAPI-positive anaphase bridges and accumulation micronuclei around bi-nucleated cells upon treatment with APH could be partially rescued when WT RNase H1 was over-expressed, suggesting a role of R-loop formation in chromosome mis-segregation. Over-production of the mutant RNase H1 did not have any significant effect on these phenotypes (Figure 13B, C).

Another cytological marker of these mitotic defects is an increase in the number of 53BP1 nuclear bodies in newly born G1 daughter cells. Upon APH treatment, about 80% of G1 cells contained more than 3 53BP1 nuclear bodies, and this was reduced when R-loops are resolved by over-expressing WT RNase H1, but not upon R-loop stabilization with the RNase H1 mutant (Figure 13D). These data suggest that under conditions of RS, cells present signs of genomic instability and defects in mitosis, and formation of R-loops is a contributing factor to the observed aberrations.

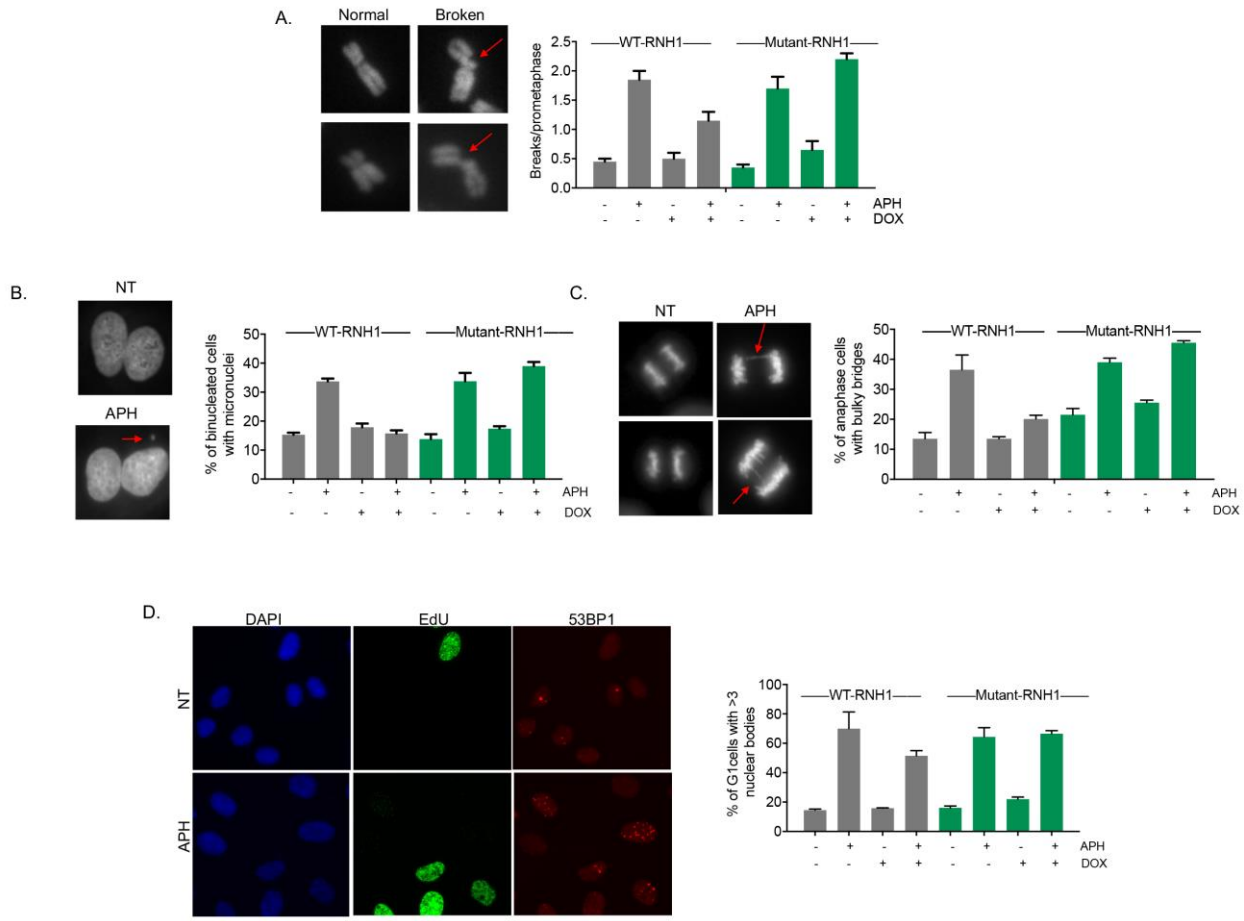


Figure 13: Contribution of R-loops to genomic instability and chromosome mis-segregation upon replication stress

A. Effect of RNase H1 (RNH1) overexpression on chromosome fragility induced by replication stress. Left panel: Examples of intact and broken prometaphase chromosomes of cells treated with APH (0.2 µM) for 16 h. Arrows denote chromatid breaks. Right panel: Quantification of chromatid breaks in U2OS T-REx cells treated with (+) or without (-) APH. Where indicated (+), expression of wild-type (WT) RNH1 (grey bars) and mutant RNH1 (D210N; green bars) was induced with 1ng/ml dox for 24 h prior to treatment. Data are represented as mean of three independent experiments; error bars show SEM.

B. Effect of RNH1 overexpression on the formation of micronuclei upon replication stress. Left panel: Representative images of bi-nucleated cells without (top) and with (bottom) micronuclei (remove NT and APH). Cells were treated with APH (0.2 µM) and Cytochalasin B (2µg/ml) for 16 h. Red arrow indicates a micronucleus. Right panel: Quantification of percentage of bi-nucleated cells with micronuclei for indicated conditions. Expression of WT RNH1 (grey bars) and Mutant RNaseH1 (green bars) was induced (+) with 1ng/ml dox for 24 h prior to treatment with APH (or DMSO) and Cytochalasin B. Data are represented as mean of three independent experiments; error bars show SEM (*right*).

C. Effect of RNH1 overexpression on the formation anaphase bridges upon replication stress. Left panel: Examples of anaphase cells with or without DAPI-positive bulky bridges (red arrows). Right panel: Quantification of the frequency of DAPI-positive anaphase bridges in cells treated with (+) or without (-) APH (0.2 µM). Expression of WT Rnase H1 (grey bars)

and mutant Rnase H1(green bars) was induced with 1ng/ml dox for 24 h prior to treatment. Data are means of two independent experiments; error bars show SEM (*right*). At least 75 anaphase cells were scored in each experiment.

D. Effect of RNH1 overexpression on accumulation of DNA damage in G1 cells upon replication stress. Left panel: Examples of G1 cells (EdU negative) containing 53BP1 nuclear bodies (red). Right panel: Quantification of G1-phase specific 53BP1 nuclear bodies in cells treated with (+) or without (-) APH (0.2 μM). Expression of WT Rnase H1 (grey bars) and mutant Rnase H1(green bars) was induced (+) with 1ng/ml dox for 24 h. Percentage of G1 cells with >3 nuclear bodies is plotted. Data are means of two independent experiments. Error bars show SD. At least 300 G1 cells were scored in each experiment.

4.5 R-loops accumulate at chromosomal fragile sites upon replication stress

Having observed R-loop formation in cells upon RS conditions that induce fragile site expression, we wondered if R-loops could be detected at these loci. To this end, we performed a series of ChIP-qPCR experiments to monitor RNH1(D210N)-GFP occupancy on chromosomal fragile sites prior to and after exposure of cells to different forms of RS. We found that upon mild RS generated by a prolonged exposure of cells to low doses of APH, RNH1(D210N)-GFP was significantly enriched on CFSs- FRA3B and FRA16D, but not on genes that are highly transcribed in early S-phase (putative ERFSs) – ACT3, APOE, RPL22, RSP19 (Figure 14A). In contrast, upon replication arrest with hydroxyurea, RNH1(D210N)-GFP was significantly enriched on genes that are highly transcribed in early S-phase but not on late-replicating CFSs; intergenic regions DHFR4 and ACT5 were used as control (Figure 14B).

These results were verified by performing DNA:RNA immunoprecipitation (DRIP) using the S9.6 antibody under the same conditions. RNA:DNA hybrids were enriched only at CFSs (FRA3B and FRA16D) upon APH treatment and specifically at actively transcribed genes (RPL13A, ACT3,APOE) when exposed to high dose of HU. Pre- treatment of immunoprecipitated samples with RNase H1 dramatically decreased the level of hybrids, confirming the signal is specific for R-loops (Figure 14C,D). From these observations, we could conclude that R-loops from at CFSs and possibly at ERFSs, under conditions of RS.

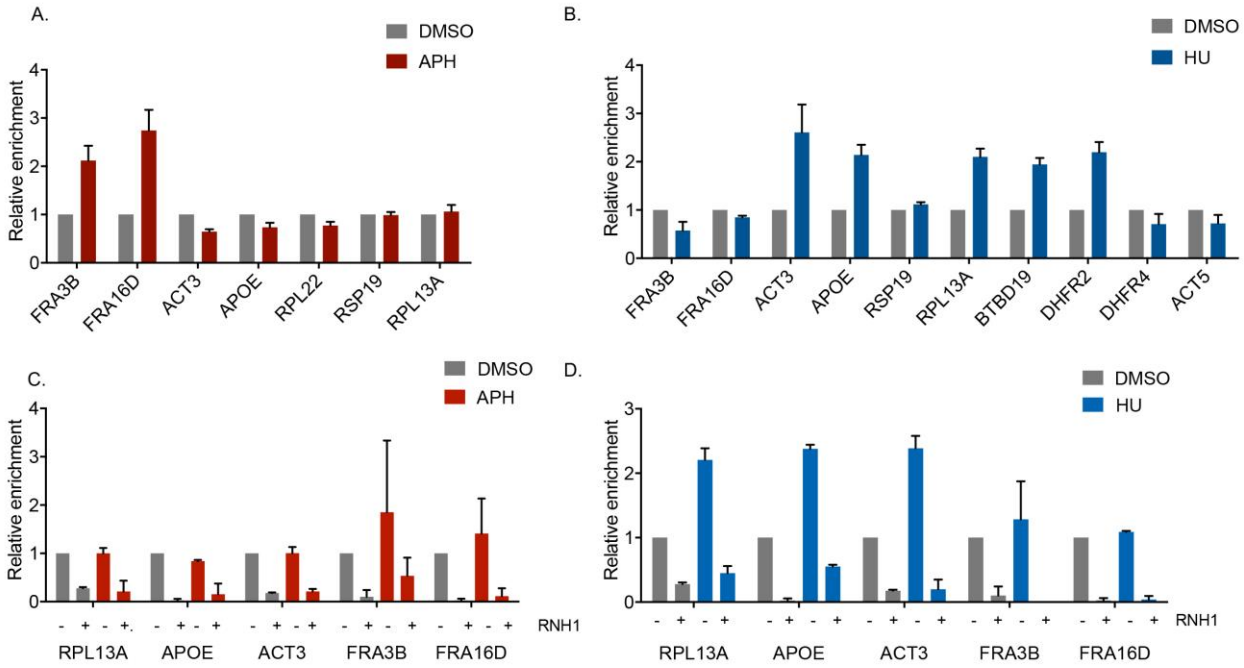


Figure 14: R-loop formation at chromosomal fragile sites

A. U2OS T-REx cells expressing RNH1(D210N)-GFP were either treated with DMSO (mock) or APH (0.2 µM) for 16hrs. After-pre-extraction, cells were fixed, chromatin was isolated and sheared, prior to immunoprecipitation with anti-GFP antibody. ChIP-qPCR was performed and enrichment was calculated relative to the input and then normalized to the mock treatment. Data are mean of three independent experiments. Error bars show SD.

B. RNH1(D210N)-GFP-expressing cells were treated with DMSO (mock) or HU (10mM) for 6hrs. ChIP-qPCR analysis was performed as in A.

C. DRIP-qPCR using S9.6 mouse monoclonal antibody. U2OS T-REx cells expressing RNH1(D210N)-GFP were treated with DMSO (mock) or APH (0.2 µM) for 16hrs. Pre-immunoprecipitated samples were untreated (-) or treated (+) with Rnase H1 (RNH1) as indicated. Signal of RNA:DNA hybrids was calculated relative to input values and normalized to mock treatment. Data represent mean ± SEM from two independent experiments.

D. RNH1(D210N)-GFP-expressing cells were treated with DMSO (mock) or HU (10 mM) for 6hrs. DRIP-qPCR was performed using S9.6 antibody as in C.

4.6 R-loops are abundant at centromeres, telomeres and rDNA regions

Since our ChIP-qPCR experiments revealed that R-loops form at distinct loci under different forms of RS, we wanted to capture a genome-wide view of R-loop formation under these conditions. To this end, we performed ChIP-sequencing to monitor RNH1(D210N)-GFP occupancy upon mock-, APH- and HU-treated conditions. An input sample was included as control. We obtained a total of around 100 million of reads per sample (~60 million uniquely mapped reads after removal of PCR duplicates). Biological replicates demonstrated a high reproducibility as observed by pair-wise comparison of independent experiments (Figure 15).

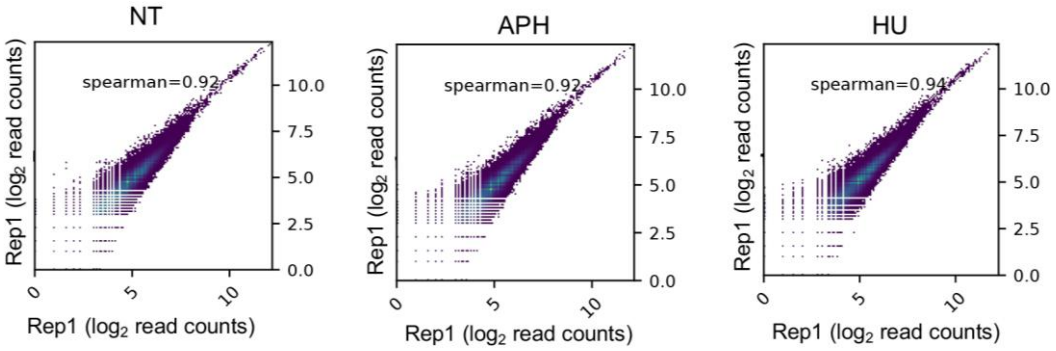


Figure 15: Reproducibility of biological replicates
Comparison of biological replicates of two ChIP-sequencing experiments performed to monitor R-loop formation by measuring RNH1(D210N)-GFP occupancy across the genome. Chromatin from U2OS T-REx cells expressing RNH1(D210N)-GFP, either treated with DMSO (NT) or APH (0.2 µM) for 16 h or HU (10 mM) for 6 h, was immunoprecipitated with anti-GFP antibody, amplified and subjected to high-throughput sequencing. Average read coverage in each continuous 3 Kb bin across the whole genome was calculated for each dataset. Spearman correlation coefficient was computed from each comparison to evaluate the reproducibility.

Genome-wide R-loop profiling in yeast revealed accumulation of R-loops at centromeres and telomeres [117]. Consistently, we could observe a significant increase in the percentage of reads at these regions as compared to the input for all the three conditions (Figure 16A, B). Interestingly, R-loops at telomeric repeats were specifically enriched upon APH treatment (Figure 16B). This is in agreement with the finding that telomeres are intrinsically difficult to replicate regions and resemble CFSs upon mild RS [133]. We also observed an increase in R-loop formation at 5S and 28S rDNA repeats, regions of high transcriptional activity, found to be associated with R-loop formation[81] (Figure 16C). Hence having detected R-loops at expected regions, we conclude that our tool can reliably detect R-loops across the genome.

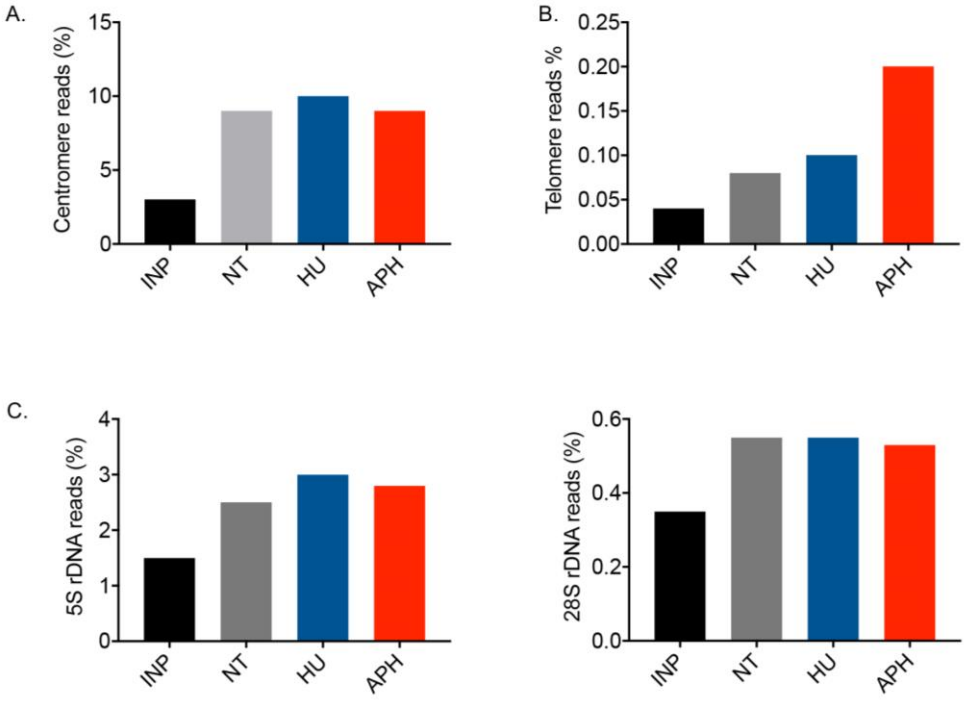


Figure 16: R-loop formation at centromeres, telomeres and rDNA
A. Percentage of reads that map to alpha-satellite centromeric repeats in input (INP), NT, APH and HU treated samples computed using deepTools.
B. Percentage of reads mapped to telomeres: defined as a stretch of 5 or more TTAGGC repeats, for INP, NT, APH and HU samples.
C. Percentage of rDNA reads: 5S rDNA *(left)* and 28S rDNA *(right)*; measured using SortMeRNA program, for reads obtained from INP, NT, APH and HU samples.

4.7 R-loops are formed at specific loci upon replication stress

Using standard peak calling methods, we identified a total of ~100,000 peaks for all treatments and biological replicates, over input as the control. To identify peaks specifically enriched upon RS, we defined significant peaks as those among the 100,000 peaks that are similar within but differ between the non-treated (NT) and treatment groups across two ChIP-Seq experiments. HU- and APH-specific peaks detected over a two-fold enrichment were considered for further analysis (Figure 17A). Upon HU treatment, 374 peaks were identified, while for APH treatment, 676 peaks were enriched over NT. Of these peaks, 207 were common between the two conditions (Figure 17B). To visualize the differential coverage of these peaks between the treatments, we generated heat maps for HU-specific and APH-specific peaks. While most of the peaks represented a higher coverage upon RS, a small fraction of peaks were more enriched prior to the treatment (NT) (Figure 17C). As an

example, we viewed one of the APH-specific peaks, obtained at chr17: 22521210-22521511, using the integrated genome viewer. We could observe an enrichment upon APH treatment as compared to NT or HU-treatments, and this was true for both experiments (Figure 17D).

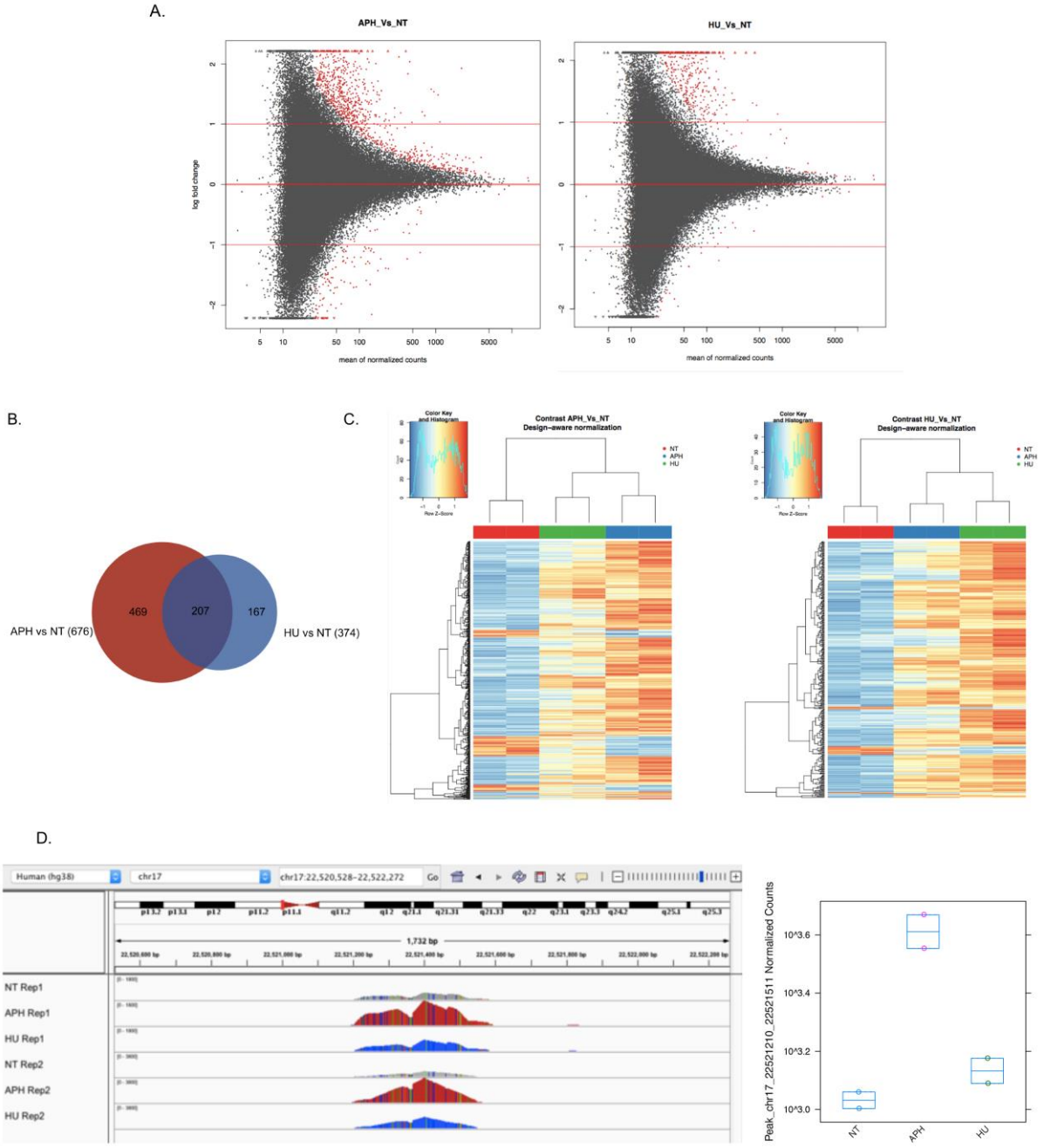


Figure 17: R-loops detected at specific loci upon replication stress

A. MA plots to visualize differences in distribution between APH-treated and NT samples *(left)* and HU-treated and NT samples (right). About 100,000 peaks (union of all peaks enriched over the input) are shown in the scatterplot. The red color marks the peaks that are differentially enriched between NT and HU or APH samples with FDR < 0.1. The peaks above the top red line are at least two-fold enriched in treated samples while ones below the bottom red line are at least two-fold enriched in NT samples.

B. Venn diagram representing the overlap between APH (676) and HU (374) peaks. 207 peaks are shared between the two treatments, while 469 peaks are APH-specific and 167 HU-specific.
C. Heat map to visualize coverage of the peaks across treatments, low z- score (blue) represents less coverage to high z –score (red) shows high coverage. *Left*: Coverage of APH-specific peaks are represented in NT, HU and APH treatments. *Right:* Coverage HU-specific peaks are represented in NT, HU and APH treatments.
D. Example of a differentially enriched peak. *Left:* Integrated Genome Viewer (IGV) display of ChIP-Seq coverage at a representative genomic region (chr17: 22521210-22521511) in NT, APH and HU samples of two biological replicates. *Right:* Box plot measuring relative changes in coverage at chr17: 22521210-22521511 across treatments.

A pre-requisite for R-loop formation is the presence of high GC content and GC-skewness: asymmetry in the distribution of G and C between the two DNA strands [50]. A G-rich non-template strand, when exposed as ssDNA in an R-loop, is prone to formation of secondary structures (G-quadraplexes) which further stabilize the R-loops [34]. When we compared the GC content of the peak sites, we could observe a significant increase compared to the expected average in the genome, for both APH- and HU-specific peaks (Figure 18A). We also measured GC-skew density around the R-loop peaks, and we could observe substantial strand asymmetry for both sets of peaks (Figure 18B).

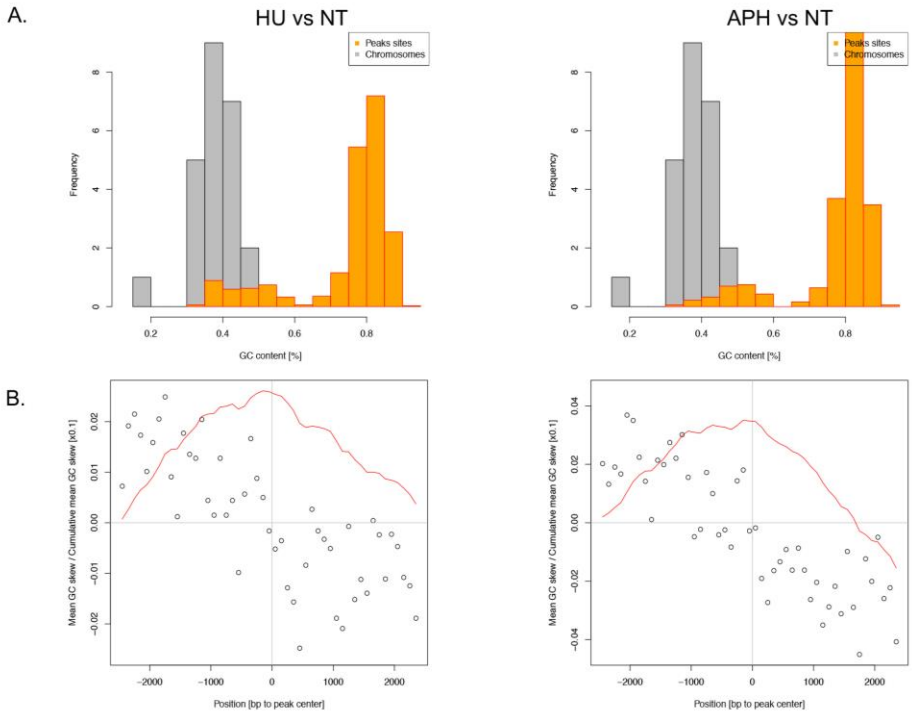


Figure 18: R-loops peaks show high GC content and GC skewness

A. GC content of peak sites (yellow) compared to GC content of complete human chromosomes (grey) calculated for HU peaks *(left)* and APH peaks *(right)* using bedTools.
B. GC skewness of R-loop peaks**:** GC skew was calculated by sliding window approach (width 100 bp) in a 5 kb region surrounding the peak center. To reduce the noise, the resulting data for all significant peaks was averaged and plotted (open circles). Cumulative GC skew is shown in red, scaled by a factor of 0.1 for HU peaks *(left)* and APH peaks *(right).*

4.8 R-loop forming loci coincide with chromosomal fragile sites upon replication stress

Having identified regions that form R-loops specifically upon RS, we wanted to further characterize these loci to decipher why they are prone to R-loop formation under these conditions. We first looked at the genomic distribution of R-loop peaks. While the R-loop forming peaks were mainly found at intergenic (~49%) and intronic (~43%) regions in unperturbed conditions, we observed that R-loops that are formed upon RS, are located mostly in the exons, introns and promoter regions (Figure 19). This difference in distribution could be a consequence of specific regions being more susceptible to transcription-replication collisions, and hence R-loop formation, under conditions of RS.

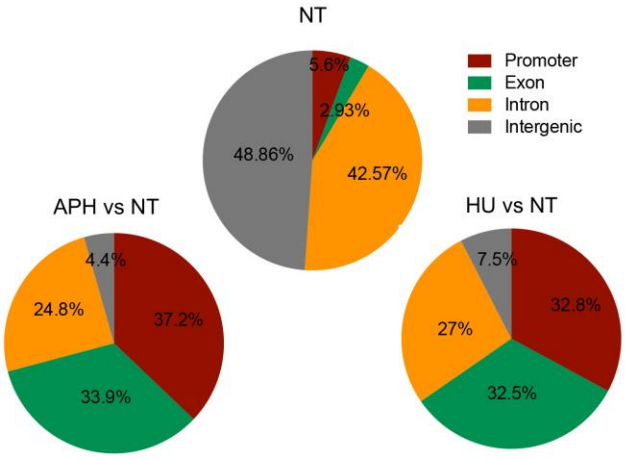


Figure 19: Distribution of R-loops in the genome
Genomic distribution of R-loop peaks across promoters, intron, exon and intergenic regions for NT, APH-specific and HU-specific peaks, computed using bedTools.

Common fragile sites have been identified as regions prone to transcription-replication conflicts when cells are exposed to RS [26]. We therefore checked the overlap of the R-loop forming loci identified with 127 fragile site regions (common fragile sites and rare fragile sites) [134]. Indeed, we could observe that a significant

fraction (~30%), of APH and HU specific R-loop peaks, coincide with fragile sites (Figure 20A). We wondered if these peaks that overlap with fragile sites are shared between APH and HU treatments. We found that of the 217 APH peaks and 121 HU peaks, that coincide with fragile site loci, only 60 peaks were common between the two, suggesting that exposure of cells to different forms of RS, can promote fragility at different regions (Figure 20B).

Fragile sites are susceptible to increased rate of DNA breakage and have been associated with deletions in cancers [135-137]. However, besides fragile site loci, most homozygous deletions in cancer genomes are unexplained. We wondered if RS-associated R-loop forming loci coincide with regions that are frequently found deleted in cancers. For this, we used results from a study that identified 2,428 somatic homozygous deletions in 746 cancer cell lines, and checked if R-loop forming regions overlapped with these deletions [134]. We identified ~55 APH-specific peaks and ~15 HU-specific peaks to be located within these recurrently deleted regions. However, this overlap is less than what is expected by chance, and hence not significant to conclude that R-loop formation is associated with deletions in cancer (Figure 20C). Further bioinformatics analysis is required to better understand the causes and consequences of R-loop formation at these specific loci upon RS.

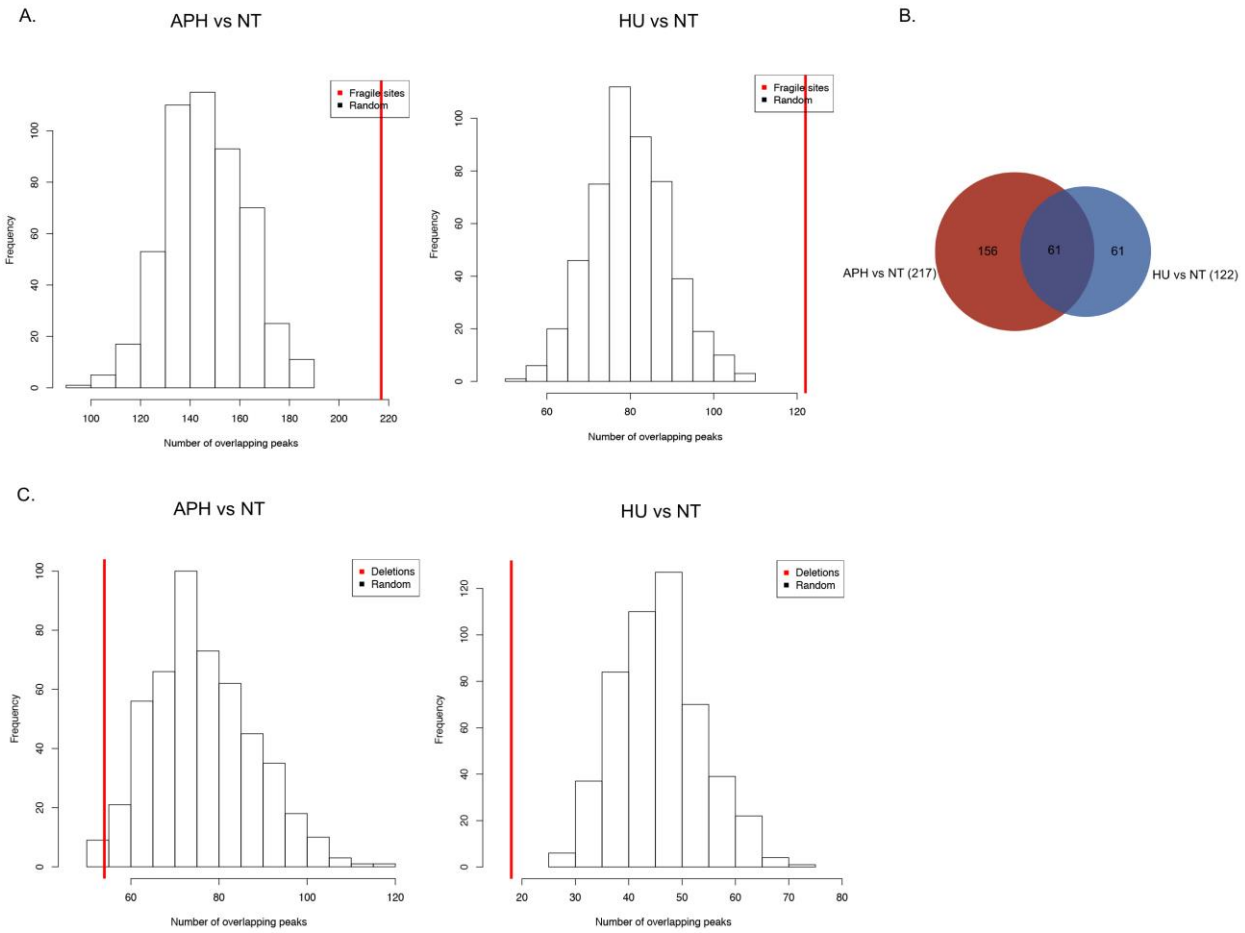


Figure 20: R-loop formation at fragile sites upon replication stress

A. Number of R-loop peaks overlapping fragile sites (represented by the red line). Histogram represents the frequency distribution of number of peaks overlapping regions obtained by random shuffling of fragile sites across the genome. *Left*: APH-specific peaks *Right*: HU-specific peaks.

B. Venn diagram representing the overlap between APH (217) and HU (121) peaks that coincide with fragile sites. 61 peaks are shared between the two treatments, while 156 and 61 peaks are unique to APH and HU treatments respectively.

C. Number of R-loop peaks overlapping homozygous deletions found in cancer (represented by the red line). Histogram represents the frequency distribution of number of peaks overlapping regions obtained by random shuffling of deletion sites across the genome. *Left*: APH-specific peaks *Right*: HU-specific peaks

5. DISCUSSION

Genomic instability arises from a variety of cellular processes in the genome, including DNA replication and transcription. Co-transcriptional R-loops result from the generation of stable RNA:DNA hybrids between the nascent transcript and the non-template DNA strand, and contribute to genomic instability. Various studies indicate that R-loops can form in cells due to defects in mRNA processing and are stabilized in the presence of a G-rich non-template strand, eventually causing replication stress (RS) by posing a roadblock to replication fork progression [36, 73, 77, 78]. However, whether RS can promote R-loop formation is not well understood. In this study, using catalytically-inactive RNase H1 as a tool to detect R-loops, we could show that R-loops form in response to different forms of RS. Moreover, we provide evidence that formation of R-loops contributes to the impairment of replication fork movement and to defects in chromosome segregation observed upon RS.

5.1 R-loops and replication stress: Causes

We could observe increased R-loop formation in cells upon different RS conditions: mild RS by APH, replication arrest by HU, Top1 inhibition by CPT, and Chk1 inhibition by UCN-01. By analyzing replication fork progression at a single molecule level using DNA fiber assay, we could show that when replication is impaired using chemical inhibitors, the observed fork slowing results not just from DNA polymerase inhibition (APH) or dNTP depletion (HU), but also from R-loop formation. This was evident from the observation that sister forks fired from normal replication origins displayed patterns of asymmetry upon induction of RS, indicating that one of the forks encountered a barrier during its progression. Inhibition of transcription or removal of R-loops by over-expressing RNase H1, completely suppressed fork asymmetry pattern, suggesting that this barrier arises from collisions between transcription and replication, and subsequent R-loop formation. In spite of a spatio-temporal segregation of transcription and replication in eukaryotes, transcription-replication conflicts are still frequent events in certain genomic regions, and it was shown that these encounters are detrimental to the cell only if replication is perturbed [26]. Based on these findings, it is tempting to speculate that while a healthy

replication fork can dislodge the transcription complex, an impaired replication fork (in the presence of RS) blocks the elongating RNA polymerase, giving sufficient time for the nascent mRNA to reanneal to template DNA to form an R-loop, which blocks DNA replication.

Shedding light on the impact of directionality on transcription-replication encounters, *Hamperl et al.* used an episomal system to show that head-on collisions promote R-loop formation while co-directional collisions resolve R-loops [24]. Co-directional encounters push RNA polymerase in the same direction as the replisome as positive DNA supercoiling ahead of the replication fork neutralizes negative supercoiling generated behind RNA polymerase. In contrast, frontal clashes may cause accumulation of positive supercoils between the two machineries, hindering both transcription and replication, and increasing R-loop formation. Although we were unable to decipher the directionality of collisions in our study, we speculate that RS increases the frequency of head-on encounters and hence elevates R-loop levels.

Why should there be an increase in head-on collision events upon replication stress? One explanation would be the activation of dormant origins. It is known that under conditions of RS, dormant origins fire locally around stalled forks, in order to complete replication [9, 138]. While a co-directional bias in the genome restricts head-on encounters between transcription and replication initiating from constitutive origins [20], dormant origin firing can disrupt this balance. In our sister fork analysis, bi-directional forks commenced from constitutive origins, as stress was induced only during the second labeling. However, we can speculate that the block we observed (asymmetry) could come from HO collisions (and R-loop formation) between a transcription complex and a replication fork from dormant origins fired in the vicinity. DNA combing experiments to measure inter-origin distance would be necessary to prove this hypothesis.

5.2 R-loops and replication stress: Consequences

Under conditions of RS, there are regions in the genome that are susceptible to breakage, and some of these have been identified as hot-spots for transcription-replication collisions, and possibly R-loop formation [26]. In our study, using ChIP-

qPCR analysis, we could detect R-loop formation at the CFSs FRA3B and FRA16D, when cells were exposed to low dose of APH. CFSs are expressed under conditions of mild RS, when cells enter mitosis with under-replicated DNA [139]. When we analyzed metaphase chromosomes after APH treatment, we could detect chromatid breaks/gaps, which were reduced upon R-loop removal by RNase H1 over-expression. Persistent R-loops in mitosis can then cause defects in chromosome segregation, and indeed we could observe that formation of anaphase bridges, micronuclei and 53BP1 nuclear bodies in G1 (signs of chromosome mis-segregation) were partially dependent on R-loop formation. These defects in mitosis can give rise to chromosomal aberrations. Of note, FHIT and WWOX, genes that house the fragile sites FRA3B and FRA16D respectively, have been classified as tumor suppressors and found deleted in multiple cancers [140-142]. Thus, formation of R-loops may contribute to these deletions and cancer development.

Besides CFSs, a new category of fragile sites- ERFSs were identified. These early replicating regions are characterized by the presence of actively transcribed gene clusters and high density of origins; and they are expressed upon replication arrest by HU [30]. In our study, we found accumulation of R-loops at actively transcribed genes when cells were exposed to high dose of HU. The increased replication initiation events near highly transcribed genes could increase conflicts between DNA replication and transcription machineries, lead to R-loop formation and thereby promote their fragility.

5.3 Genome-wide approach: Bird's eye view of R-loop formation

To obtain a more complete picture about R-loop formation upon RS, we performed a genome-wide analysis of binding sites of catalytically-inactive RNase H1 in unperturbed (NT, non-treated), APH- and HU-treated cells. Under all the three conditions, we observed accumulation of R-loops at repetitive sequences like centromeres and telomeres, and at highly transcribed rDNA repeats. The percentage of reads at telomeres was significantly higher in APH-treated cells. In agreement with this finding, earlier studies have shown that these repetitive telomeric stretches pose a challenge to the DNA replication machinery, giving rise to replication-dependent defects that resemble those of CFSs as exposing cells to low doses of APH

increased telomere fragility [133]. Our data suggests that R-loop formation may contribute to this observed telomere fragility. Measuring telomeric FISH signal at individual chromatid ends of metaphase chromosomes upon RS while over-expressing RNase H1 would help confirm this theory.

The main aim of our ChIP-seq experiments was to identify the loci that form R-loops specifically under conditions of RS. We observed 674 peaks enriched upon APH treatment and 375 peaks upon HU treatment. Between the two RS conditions, 120 peaks overlapped with each other. We compared the genomic distribution of specific R-loop peaks after RS (APH and HU) to that of the total R-loop peaks obtained. While prior to RS, a large fraction of R-loops form in intergenic regions and introns, upon RS, R-loop formation is significantly higher in exons and promoter regions. Non-coding regions of the genome (introns and intergenic) serve as a mutational buffer in eukaryotic genomes, protecting coding sequences from being affected by randomly occurring deleterious mutations [143]. Since R-loop formation can cause mutations, cells may maintain physiological R-loops to form within these regions, to preserve genome integrity. However, upon RS, formation of additional R-loops can occur in exons as well as introns, but less so in the intergenic regions; owing to an increase in transcription-replication conflicts within genes.

Genome-wide studies in different cell lines have reported that under unperturbed conditions, R-loops are commonly found at promoters of genes, where they may restrict promoter methylation and thus promote transcription of these genes [48, 49]. An intriguing question would be if the formation of R-loops at promoters upon RS modulates the expression level of genes, and if so, what role do these proteins have in RS conditions. For example, we could detect R-loops at the promoters of oncogene Cyclin E (CCNE1) and JunD proto-oncogene (JUND), after APH and HU treatments; up-regulated expression of these genes may contribute to genome instability. However, R-loop formation is not always detrimental to the cells, they play a role in many physiological processes (Discussed in Introduction section). It would be interesting to speculate if R-loop formation we observed at promoters of some DNA repair genes like ATM interacting protein (ATMIN), and Tonsoku-like protein (TONSL), upon RS, would increase the expression of these genes as a mechanism to deal with the stress. Validation of these theories would require RNA-

sequencing to be performed after APH and HU treatments, to find the correlation between R-loop levels at promoters and gene expression.

To further understand the consequence of R-loop formation upon RS, we checked if these regions overlap with chromosomal fragile sites or regions recurrently found deleted in cancers. Fragile sites, regions susceptible to breakage under mild RS, are classified as rare or common, depending on their frequency within the population. Rare fragile sites appear in the chromosomes of only a small fraction of the human population while CFSs are present in all individuals and are thus considered to be an intrinsic part of the chromosomal structure [144]. A total of 127 fragile sites (common and rare) have been characterized, and we found 217 of APH-specific R-loop-forming loci overlapped with these sites, suggesting formation of R-loops may contribute to their fragility. We also studied R-loop formation under conditions of replication arrest (HU), which according to a recent study, induces the expression of another category of fragile sites- ERFSs, but does not induce CFS expression [30]. Since these sites were identified in mouse lymphocytes, and have not been characterized in human cells, we could not check if HU-specific R-loop peaks overlap with these loci. However, to our surprise, we found that significant number HU-specific peaks overlapped with common and rare fragile sites.

Many fragile site loci have been found deleted in cancers [135-137]. We wondered if R-loop formation at these regions could be the cause of the underlying deletions. To check if there might be a link between R-loop formation and deletions found in cancers, we asked if R-loop forming regions specific to RS conditions overlapped with homozygous deletions (obtained from a study that screened 746 cancer cell lines for deletions) [134]. We did not observe a significant co-relation between these deletions and R-loop formation in this analysis. We would like to verify this by using other available data sets. For example, a recent study identified certain DNA sequences with the potential to fold into secondary structures [potential non-B DNA structures (PONDS)], such as triplexes, quadruplexes, hairpin/cruciforms, Z-DNA and single-stranded looped-out structures, at or in close proximity to the rearrangement breakpoints (translocations and deletion breakpoints) found in cancer [145]. Since R-loop forming sequences may fall in the category of PONDS, it would be interesting to overlap our data with the rearrangement breakpoints used in this study, to understand if R-loop formation has a role in cancer.

5.4 RNase H1 as an R-loop detection tool

RNase H1 is a nuclease that cleaves the RNA strand in RNA:DNA hybrids, and its over-expression has been used to rescue phenotypes observed upon R-loop formation. In our study, we used a GFP-tagged, catalytically-inactive mutant of RNase H1, which can bind to, but not cleave the hybrid. Although novel at the start of this project, some recently published studies have also employed this tool. DRIVE-seq, (DNA:RNA in-vitro enrichment) uses a catalytically inactive MBP-RNase H1, to enable affinity pull-down of R-loop forming regions [48]. GFP-tagged versions of nuclease-deficient RNase H1 have been used to visualize R-loops by fluorescence microscopy [59, 70, 85]. A very recent study developed a method called R-ChIP, where catalytically-inactive RNase H1 is used to capture R-loop forming genomic regions by chromatin immunoprecipitation followed by genome-wide sequencing, enabling mapping R-loops with a higher resolution than already available methods [146].

The most abundantly used tool to study R-loops is the monoclonal antibody (S9.6) developed against RNA:DNA hybrids. However, some inconsistencies in R-loop detection owing to its specificity have been reported. For instance, a recent study showed that S9.6 antibody could detect R-loops with a bias for certain sequences within the hybrid [147]. Another study reported that the affinity of the antibody for dsRNA results in inaccurate R-loop mapping [123]. Genome-wide studies with the S9.6 antibody also suffer from poor resolution due to the relatively large size of restriction fragments used for immunoprecipitation. In order to improve the resolution, isolated DNA was sonicated prior to antibody capture, but this could damage R-loops, compromising the data quality [119]. Another modification of the technique was DRIPc-seq, which involved sequencing associated RNA, however any residual RNA tightly bound to chromatin will give a high background signal [148]. Our tool specifically recognizes R-loops *in vivo* under different biological conditions, and we used standard ChIP-seq protocol, which involved fixation to stabilize R-loops and sonication to increase the resolution, providing us with high quality data. The drawback of this system is that its use is restricted to a single cell line. Developing a system wherein inactive RNase H1 can be introduced into different cell lines would help broaden its applications.

5.5 Conclusion

We established our own tool for the detection and isolation of R-loops generated in U-2 OS cells. With the help of this tool, we could show that R-loops are formed in response to different forms of RS. We propose a model wherein we hypothesize that under normal conditions, transcription encounters replication forks co-directionally, owing to the co-directional bias in the genome, avoiding R-loop formation and the subsequent blockage of replication fork progression (Figure 21). If some head-on encounters occur, unperturbed forks can dislodge RNA polymerase and resolve any possible conflicts. Cells therefore enter mitosis without any problems and give rise to healthy daughter cells, preserving genome integrity. However, under conditions of RS, firing of dormant origins enhances the probability of transcription meeting impaired replication forks in the HO orientation (Figure 21). These slow forks are unable to displace RNA polymerase, stalling both the replisome and transcription complex, which provides sufficient time for the nascent mRNA to reanneal to the template DNA, giving rise to stable R-loops. These collision complexes block fork progression not just at the site of collision, but also serve as a roadblock to other replication forks originating from constitutive origins (Figure 21). Cells entering mitosis with unresolved conflicts have impaired chromosome segregation, leading to genomic instability (Figure 21).

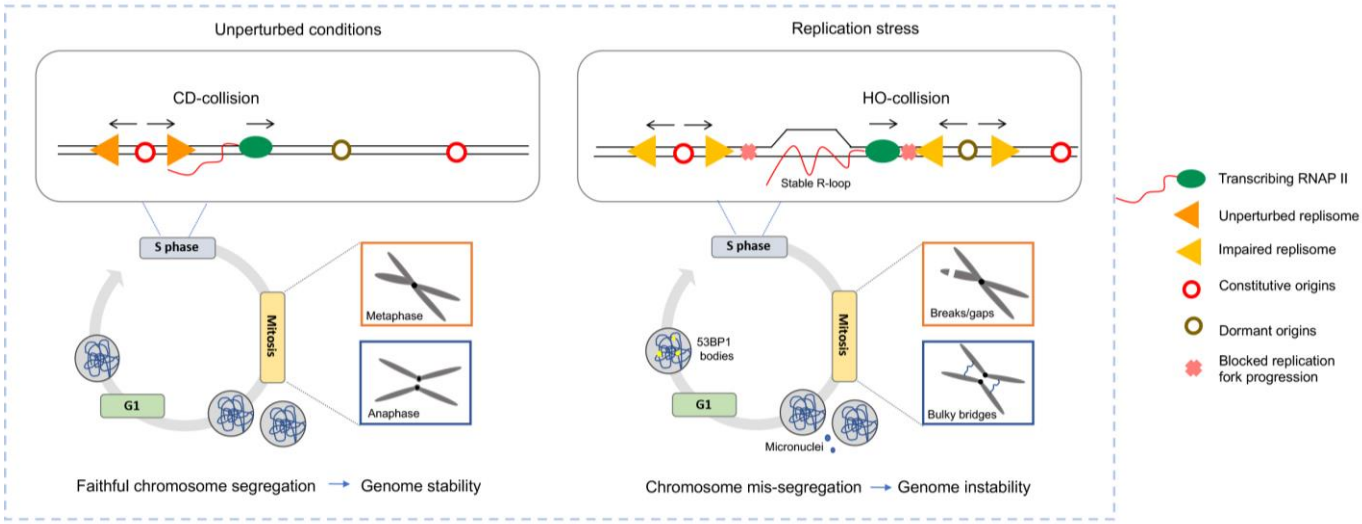


Figure 21: Proposed model for R-loop formation upon replication stress
For detailed description see the section 5.5.

5.6 Relevance of the study

Evidence indicates that oncogene activation in pre-cancerous lesions gives rise to replication stress and DNA damage [149]. A major consequence of oncogene activation is altered origin usage (Discussed in Introduction section). Increase in origin firing can mediate replication stress by causing depletion of dNTPs or slow-down of DNA polymerase due to lack of replisome components. In our study, we mimicked these situations by the use of HU and APH, respectively. Hence, we can speculate that oncogene-induced replication stress in pre-cancerous cells can also lead to transcription-replication conflicts and R-loop formation. Measuring the level of R-loops in patient samples could help identify pre-cancerous lesions. Thus R-loops could be a novel diagnostic biomarker for cancers.

Genome-wide analysis of R-loop formation upon RS helped us identify unique R-loop forming sites upon APH and HU treatments. Although we obtained distinct peaks between the two conditions, a fraction of the peaks overlap. Performing genome-wide analysis of R-loop formation upon RS induced by activation of different oncogenes and drugs, can help create a database of common R-loop forming loci under all these conditions. This information can help in the screening of pre-cancerous lesions based on their R-loop profiles.

6. BIBLIOGRAPHY

1. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
2. Negrini, S., V.G. Gorgoulis, and T.D. Halazonetis, *Genomic instability--an evolving hallmark of cancer.* Nat Rev Mol Cell Biol, 2010. **11**(3): p. 220-8.
3. Gorgoulis, V.G., et al., *Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions.* Nature, 2005. **434**(7035): p. 907-13.
4. Bartkova, J., et al., *DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis.* Nature, 2005. **434**(7035): p. 864-70.
5. Neelsen, K.J., et al., *Oncogenes induce genotoxic stress by mitotic processing of unusual replication intermediates.* J Cell Biol, 2013. **200**(6): p. 699-708.
6. Halazonetis, T.D., V.G. Gorgoulis, and J. Bartek, *An oncogene-induced DNA damage model for cancer development.* Science, 2008. **319**(5868): p. 1352-5.
7. Remus, D., et al., *Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing.* Cell, 2009. **139**(4): p. 719-30.
8. Evrin, C., et al., *A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication.* Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20240-5.
9. Woodward, A.M., et al., *Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress.* J Cell Biol, 2006. **173**(5): p. 673-83.
10. Burrell, R.A., et al., *Replication stress links structural and numerical cancer chromosomal instability.* Nature, 2013. **494**(7438): p. 492-496.
11. Kawabata, T., et al., *Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression.* Mol Cell, 2011. **41**(5): p. 543-53.
12. Ekholm-Reed, S., et al., *Deregulation of cyclin E in human cells interferes with prereplication complex assembly.* J Cell Biol, 2004. **165**(6): p. 789-800.
13. Jones, R.M., et al., *Increased replication initiation and conflicts with transcription underlie Cyclin E-induced replication stress.* Oncogene, 2013. **32**(32): p. 3744-53.
14. Dominguez-Sola, D., et al., *Non-transcriptional control of DNA replication by c-Myc.* Nature, 2007. **448**(7152): p. 445-51.
15. Macheret, M. and T.D. Halazonetis, *Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress.* Nature, 2018. **555**(7694): p. 112-116.
16. Mejlvang, J., et al., *New histone supply regulates replication fork speed and PCNA unloading.* J Cell Biol, 2014. **204**(1): p. 29-43.
17. Toledo, L.I., et al., *ATR prohibits replication catastrophe by preventing global exhaustion of RPA.* Cell, 2013. **155**(5): p. 1088-103.
18. Hangauer, M.J., I.W. Vaughn, and M.T. McManus, *Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs.* PLoS Genet, 2013. **9**(6): p. e1003569.
19. Bermejo, R., M.S. Lai, and M. Foiani, *Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription.* Mol Cell, 2012. **45**(6): p. 710-8.

20. Petryk, N., et al., *Replication landscape of the human genome.* Nat Commun, 2016. **7**: p. 10208.
21. Kotsantis, P., et al., *Increased global transcription activity as a mechanism of replication stress in cancer.* Nat Commun, 2016. **7**: p. 13087.
22. Gros, J., et al., *Post-licensing Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA.* Mol Cell, 2015. **60**(5): p. 797-807.
23. Powell, S.K., et al., *Dynamic loading and redistribution of the Mcm2-7 helicase complex through the cell cycle.* EMBO J, 2015. **34**(4): p. 531-43.
24. Hamperl, S., et al., *Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses.* Cell, 2017. **170**(4): p. 774-786 e19.
25. Ozeri-Galai, E., A.C. Bester, and B. Kerem, *The complex basis underlying common fragile site instability in cancer.* Trends Genet, 2012. **28**(6): p. 295-302.
26. Helmrich, A., M. Ballarino, and L. Tora, *Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes.* Mol Cell, 2011. **44**(6): p. 966-77.
27. Di Marco, S., et al., *RECQ5 Helicase Cooperates with MUS81 Endonuclease in Processing Stalled Replication Forks at Common Fragile Sites during Mitosis.* Mol Cell, 2017. **66**(5): p. 658-671 e8.
28. Naim, V., et al., *ERCC1 and MUS81-EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis.* Nat Cell Biol, 2013. **15**(8): p. 1008-15.
29. Ying, S., et al., *MUS81 promotes common fragile site expression.* Nat Cell Biol, 2013. **15**(8): p. 1001-7.
30. Barlow, J.H., et al., *Identification of early replicating fragile sites that contribute to genome instability.* Cell, 2013. **152**(3): p. 620-32.
31. Tsantoulis, P.K., et al., *Oncogene-induced replication stress preferentially targets common fragile sites in preneoplastic lesions. A genome-wide study.* Oncogene, 2008. **27**(23): p. 3256-64.
32. Thomas, M., R.L. White, and R.W. Davis, *Hybridization of RNA to double-stranded DNA: formation of R-loops.* Proc Natl Acad Sci U S A, 1976. **73**(7): p. 2294-8.
33. Drolet, M., et al., *Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I.* Proc Natl Acad Sci U S A, 1995. **92**(8): p. 3526-30.
34. Hamperl, S. and K.A. Cimprich, *The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability.* DNA Repair (Amst), 2014. **19**: p. 84-94.
35. Westover, K.D., D.A. Bushnell, and R.D. Kornberg, *Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center.* Cell, 2004. **119**(4): p. 481-9.
36. Roy, D. and M.R. Lieber, *G clustering is important for the initiation of transcription-induced R-loops in vitro, whereas high G density without clustering is sufficient thereafter.* Mol Cell Biol, 2009. **29**(11): p. 3124-33.
37. Wahba, L., S.K. Gore, and D. Koshland, *The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability.* Elife, 2013. **2**: p. e00505.

38. Skourti-Stathaki, K. and N.J. Proudfoot, *A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression.* Genes Dev, 2014. **28**(13): p. 1384-96.
39. Costantino, L. and D. Koshland, *The Yin and Yang of R-loop biology.* Curr Opin Cell Biol, 2015. **34**: p. 39-45.
40. Ramiro, A.R., et al., *Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand.* Nat Immunol, 2003. **4**(5): p. 452-6.
41. Chaudhuri, J., et al., *Transcription-targeted DNA deamination by the AID antibody diversification enzyme.* Nature, 2003. **422**(6933): p. 726-30.
42. Zhang, Z.Z., et al., *The strength of an Ig switch region is determined by its ability to drive R loop formation and its number of WGCW sites.* Cell Rep, 2014. **8**(2): p. 557-69.
43. Wiedemann, E.M., M. Peycheva, and R. Pavri, *DNA Replication Origins in Immunoglobulin Switch Regions Regulate Class Switch Recombination in an R-Loop-Dependent Manner.* Cell Rep, 2016. **17**(11): p. 2927-2942.
44. Guibert, S., T. Forne, and M. Weber, *Dynamic regulation of DNA methylation during mammalian development.* Epigenomics, 2009. **1**(1): p. 81-98.
45. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nat Rev Genet, 2002. **3**(6): p. 415-28.
46. Illingworth, R.S. and A.P. Bird, *CpG islands--'a rough guide'.* FEBS Lett, 2009. **583**(11): p. 1713-20.
47. Takeshima, H., et al., *The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands.* Genome Res, 2009. **19**(11): p. 1974-82.
48. Ginno, P.A., et al., *R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters.* Mol Cell, 2012. **45**(6): p. 814-25.
49. Grunseich, C., et al., *Senataxin Mutation Reveals How R-Loops Promote Transcription by Blocking DNA Methylation at Gene Promoters.* Mol Cell, 2018. **69**(3): p. 426-437 e7.
50. Ginno, P.A., et al., *GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination.* Genome Res, 2013. **23**(10): p. 1590-600.
51. Skourti-Stathaki, K., N.J. Proudfoot, and N. Gromak, *Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination.* Mol Cell, 2011. **42**(6): p. 794-805.
52. Skourti-Stathaki, K., K. Kamieniarz-Gdula, and N.J. Proudfoot, *R-loops induce repressive chromatin marks over mammalian gene terminators.* Nature, 2014. **516**(7531): p. 436-9.
53. Kabeche, L., et al., *A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation.* Science, 2018. **359**(6371): p. 108-114.
54. Mimitou, E.P. and L.S. Symington, *Nucleases and helicases take center stage in homologous recombination.* Trends Biochem Sci, 2009. **34**(5): p. 264-72.
55. Mimitou, E.P. and L.S. Symington, *DNA end resection: many nucleases make light work.* DNA Repair (Amst), 2009. **8**(9): p. 983-95.
56. Blackwood, J.K., et al., *End-resection at DNA double-strand breaks in the three domains of life.* Biochem Soc Trans, 2013. **41**(1): p. 314-20.
57. Davies, A.A., et al., *Role of BRCA2 in control of the RAD51 recombination and DNA repair protein.* Mol Cell, 2001. **7**(2): p. 273-82.

58. Baumann, P. and S.C. West, *Role of the human RAD51 protein in homologous recombination and double-stranded-break repair.* Trends Biochem Sci, 1998. **23**(7): p. 247-51.
59. Britton, S., et al., *DNA damage triggers SAF-A and RNA biogenesis factors exclusion from chromatin coupled to R-loops removal.* Nucleic Acids Res, 2014. **42**(14): p. 9047-62.
60. Ohle, C., et al., *Transient RNA-DNA Hybrids Are Required for Efficient Double-Strand Break Repair.* Cell, 2016. **167**(4): p. 1001-1013 e7.
61. Aymard, F., et al., *Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks.* Nat Struct Mol Biol, 2014. **21**(4): p. 366-74.
62. Tresini, M., et al., *The core spliceosome as target and effector of non-canonical ATM signalling.* Nature, 2015. **523**(7558): p. 53-8.
63. Su, X.A. and C.H. Freudenreich, *Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2017. **114**(40): p. E8392-E8401.
64. Sollier, J., et al., *Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability.* Mol Cell, 2014. **56**(6): p. 777-85.
65. Castellano-Pozo, M., et al., *R loops are linked to histone H3 S10 phosphorylation and chromatin condensation.* Mol Cell, 2013. **52**(4): p. 583-90.
66. Garcia-Pichardo, D., et al., *Histone Mutants Separate R Loop Formation from Genome Instability Induction.* Mol Cell, 2017. **66**(5): p. 597-609 e5.
67. Wanrooij, P.H., et al., *A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop.* Nucleic Acids Res, 2012. **40**(20): p. 10334-44.
68. Stork, C.T., et al., *Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage.* Elife, 2016. **5**.
69. Hatchi, E., et al., *BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair.* Mol Cell, 2015. **57**(4): p. 636-47.
70. Bhatia, V., et al., *BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2.* Nature, 2014. **511**(7509): p. 362-5.
71. Stirling, P.C., et al., *R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants.* Genes Dev, 2012. **26**(2): p. 163-75.
72. Huertas, P. and A. Aguilera, *Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination.* Mol Cell, 2003. **12**(3): p. 711-21.
73. Dominguez-Sanchez, M.S., et al., *Genome instability and transcription elongation impairment in human cells depleted of THO/TREX.* PLoS Genet, 2011. **7**(12): p. e1002386.
74. Salas-Armenteros, I., et al., *Human THO-Sin3A interaction reveals new mechanisms to prevent R-loops that cause genome instability.* EMBO J, 2017. **36**(23): p. 3532-3547.
75. Garcia-Benitez, F., H. Gaillard, and A. Aguilera, *Physical proximity of chromatin to nuclear pores prevents harmful R loop accumulation contributing to maintain genome stability.* Proc Natl Acad Sci U S A, 2017. **114**(41): p. 10942-10947.
76. Blobel, G., *Gene gating: a hypothesis.* Proc Natl Acad Sci U S A, 1985. **82**(24): p. 8527-9.
77. Li, X. and J.L. Manley, *Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability.* Cell, 2005. **122**(3): p. 365-78.

78. Santos-Pereira, J.M., et al., *The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability.* Genes Dev, 2013. **27**(22): p. 2445-58.
79. Paulsen, R.D., et al., *A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability.* Mol Cell, 2009. **35**(2): p. 228-39.
80. Tuduri, S., et al., *Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription.* Nat Cell Biol, 2009. **11**(11): p. 1315-24.
81. El Hage, A., et al., *Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis.* Genes Dev, 2010. **24**(14): p. 1546-58.
82. Yang, Y., et al., *Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation.* Mol Cell, 2014. **53**(3): p. 484-97.
83. Marechal, A. and L. Zou, *RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response.* Cell Res, 2015. **25**(1): p. 9-23.
84. Sikorski, T.W., et al., *Sub1 and RPA associate with RNA polymerase II at different stages of transcription.* Mol Cell, 2011. **44**(3): p. 397-409.
85. Nguyen, H.D., et al., *Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1.* Mol Cell, 2017. **65**(5): p. 832-847 e4.
86. Shin, J.H. and Z. Kelman, *The replicative helicases of bacteria, archaea, and eukarya can unwind RNA-DNA hybrid substrates.* J Biol Chem, 2006. **281**(37): p. 26914-21.
87. Vijayraghavan, S., F.L. Tsai, and A. Schwacha, *A Checkpoint-Related Function of the MCM Replicative Helicase Is Required to Avert Accumulation of RNA:DNA Hybrids during S-phase and Ensuing DSBs during G2/M.* PLoS Genet, 2016. **12**(8): p. e1006277.
88. Tutt, A. and A. Ashworth, *The relationship between the roles of BRCA genes in DNA repair and cancer predisposition.* Trends Mol Med, 2002. **8**(12): p. 571-6.
89. Schwab, R.A., et al., *The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription.* Mol Cell, 2015. **60**(3): p. 351-61.
90. Garcia-Rubio, M.L., et al., *The Fanconi Anemia Pathway Protects Genome Integrity from R-loops.* PLoS Genet, 2015. **11**(11): p. e1005674.
91. Madireddy, A., et al., *FANCD2 Facilitates Replication through Common Fragile Sites.* Mol Cell, 2016. **64**(2): p. 388-404.
92. Okamoto, Y., et al., *Replication stress induces accumulation of FANCD2 at central region of large fragile genes.* Nucleic Acids Res, 2018.
93. Belotserkovskaya, R., et al., *FACT facilitates transcription-dependent nucleosome alteration.* Science, 2003. **301**(5636): p. 1090-3.
94. Abe, T., et al., *The histone chaperone facilitates chromatin transcription (FACT) protein maintains normal replication fork rates.* J Biol Chem, 2011. **286**(35): p. 30504-12.
95. Herrera-Moyano, E., et al., *The yeast and human FACT chromatin-reorganizing complexes solve R-loop-mediated transcription-replication conflicts.* Genes Dev, 2014. **28**(7): p. 735-48.
96. Zeller, P., et al., *Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability.* Nat Genet, 2016. **48**(11): p. 1385-1395.
97. Cerritelli, S.M. and R.J. Crouch, *Ribonuclease H: the enzymes in eukaryotes.* FEBS J, 2009. **276**(6): p. 1494-505.
98. O'Connell, K., S. Jinks-Robertson, and T.D. Petes, *Elevated Genome-Wide Instability in Yeast Mutants Lacking RNase H Activity.* Genetics, 2015. **201**(3): p. 963-75.

99. Parajuli, S., et al., *Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork.* J Biol Chem, 2017. **292**(37): p. 15216-15224.
100. Crow, Y.J., et al., *Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection.* Nat Genet, 2006. **38**(8): p. 910-6.
101. Steinmetz, E.J., et al., *Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase.* Mol Cell, 2006. **24**(5): p. 735-46.
102. Mischo, H.E., et al., *Yeast Sen1 helicase protects the genome from transcription-associated instability.* Mol Cell, 2011. **41**(1): p. 21-32.
103. Alzu, A., et al., *Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes.* Cell, 2012. **151**(4): p. 835-46.
104. Yuce, O. and S.C. West, *Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response.* Mol Cell Biol, 2013. **33**(2): p. 406-17.
105. Brustel, J., et al., *Large XPF-dependent deletions following misrepair of a DNA double strand break are prevented by the RNA:DNA helicase Senataxin.* Sci Rep, 2018. **8**(1): p. 3850.
106. Cohen, S., et al., *Senataxin resolves RNA:DNA hybrids forming at DNA double-strand breaks to prevent translocations.* Nat Commun, 2018. **9**(1): p. 533.
107. Boule, J.B. and V.A. Zakian, *The yeast Pif1p DNA helicase preferentially unwinds RNA DNA substrates.* Nucleic Acids Res, 2007. **35**(17): p. 5809-18.
108. Boule, J.B. and V.A. Zakian, *Roles of Pif1-like helicases in the maintenance of genomic stability.* Nucleic Acids Res, 2006. **34**(15): p. 4147-53.
109. Zhou, R., et al., *Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA.* Elife, 2014. **3**: p. e02190.
110. Hodroj, D., et al., *An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism.* EMBO J, 2017. **36**(9): p. 1182-1198.
111. Song, C., et al., *SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability.* Genes Dev, 2017.
112. Li, L., et al., *DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks.* Mol Cell Biol, 2016. **36**(22): p. 2794-2810.
113. Li, L., E.A. Monckton, and R. Godbout, *A role for DEAD box 1 at DNA double-strand breaks.* Mol Cell Biol, 2008. **28**(20): p. 6413-25.
114. Popuri, V., et al., *The Human RecQ helicases, BLM and RECQ1, display distinct DNA substrate specificities.* J Biol Chem, 2008. **283**(26): p. 17766-76.
115. Chang, E.Y., et al., *RECQ-like helicases Sgs1 and BLM regulate R-loop-associated genome instability.* J Cell Biol, 2017. **216**(12): p. 3991-4005.
116. Boguslawski, S.J., et al., *Characterization of monoclonal antibody to DNA.RNA and its application to immunodetection of hybrids.* J Immunol Methods, 1986. **89**(1): p. 123-30.
117. Chan, Y.A., et al., *Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip.* PLoS Genet, 2014. **10**(4): p. e1004288.
118. El Hage, A., et al., *Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria.* PLoS Genet, 2014. **10**(10): p. e1004716.

119. Nadel, J., et al., *RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships.* Epigenetics Chromatin, 2015. **8**: p. 46.
120. Bonnet, A., et al., *Introns Protect Eukaryotic Genomes from Transcription-Associated Genetic Instability.* Mol Cell, 2017. **67**(4): p. 608-621 e6.
121. Dumelie, J.G. and S.R. Jaffrey, *Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq.* Elife, 2017. **6**.
122. Jenjaroenpun, P., et al., *R-loopDB: a database for R-loop forming sequences (RLFS) and R-loops.* Nucleic Acids Res, 2017. **45**(D1): p. D119-D127.
123. Hartono, S.R., et al., *The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast.* J Mol Biol, 2017.
124. Wu, H., W.F. Lima, and S.T. Crooke, *Investigating the structure of human RNase H1 by site-directed mutagenesis.* J Biol Chem, 2001. **276**(26): p. 23547-53.
125. Ghodgaonkar, M.M., et al., *Phenotypic characterization of missense polymerase-delta mutations using an inducible protein-replacement system.* Nat Commun, 2014. **5**: p. 4990.
126. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
127. Wu, T.D., et al., *GMAP and GSNAP for Genomic Sequence Alignment: Enhancements to Speed, Accuracy, and Functionality.* Methods Mol Biol, 2016. **1418**: p. 283-334.
128. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
129. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
130. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-2.
131. Kopylova, E., L. Noe, and H. Touzet, *SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.* Bioinformatics, 2012. **28**(24): p. 3211-7.
132. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.
133. Sfeir, A., et al., *Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication.* Cell, 2009. **138**(1): p. 90-103.
134. Bignell, G.R., et al., *Signatures of mutation and selection in the cancer genome.* Nature, 2010. **463**(7283): p. 893-8.
135. Hirst, W.J., B. Czepulkowski, and G.J. Mufti, *Consistent interstitial chromosomal deletions in myeloid malignancies and their correlation with fragile sites.* Cancer Genet Cytogenet, 1993. **65**(1): p. 51-7.
136. Lai, L.A., et al., *Deletion at fragile sites is a common and early event in Barrett's esophagus.* Mol Cancer Res, 2010. **8**(8): p. 1084-94.
137. Le Tallec, B., et al., *Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes.* Cell Rep, 2013. **4**(3): p. 420-8.
138. Ibarra, A., E. Schwob, and J. Mendez, *Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 8956-61.
139. Mankouri, H.W., D. Huttner, and I.D. Hickson, *How unfinished business from S-phase affects mitosis and beyond.* EMBO J, 2013. **32**(20): p. 2661-71.

140. Bednarek, A.K., et al., *WWOX, a novel WW domain-containing protein mapping to human chromosome 16q23.3-24.1, a region frequently affected in breast cancer.* Cancer Res, 2000. **60**(8): p. 2140-5.
141. Ohta, M., et al., *The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers.* Cell, 1996. **84**(4): p. 587-97.
142. Roy, D., et al., *Tumor suppressor genes FHIT and WWOX are deleted in primary effusion lymphoma (PEL) cell lines.* Blood, 2011. **118**(7): p. e32-9.
143. Jo, B.S. and S.S. Choi, *Introns: The Functional Benefits of Introns in Genomes.* Genomics Inform, 2015. **13**(4): p. 112-8.
144. Schwartz, M., E. Zlotorynski, and B. Kerem, *The molecular basis of common and rare fragile sites.* Cancer Lett, 2006. **232**(1): p. 13-26.
145. Bacolla, A., et al., *Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequences.* Nucleic Acids Res, 2016. **44**(12): p. 5673-88.
146. Chen, L., et al., *R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters.* Mol Cell, 2017. **68**(4): p. 745-757 e5.
147. Konig, F., T. Schubert, and G. Langst, *The monoclonal S9.6 antibody exhibits highly variable binding affinities towards different R-loop sequences.* PLoS One, 2017. **12**(6): p. e0178875.
148. Sanz, L.A., et al., *Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals.* Mol Cell, 2016. **63**(1): p. 167-78.
149. Bartkova, J., et al., *Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints.* Nature, 2006. **444**(7119): p. 633-7.

7. APPENDIX

7.1 RECQ5 helicase cooperates with MUS81 endonuclease in processing of stalled replication forks at common fragile sites during mitosis

Stefano Di Marco[1,+], Zdenka Hasanova[2,+], Radhakrishnan Kanagaraj[1,+,#], Nagaraja Chapiddi[1,+], Veronika Altmannova[2,3,+], **Shruti Menon[1]**, Hana Sedlackova[2], Jana Langhoff[1], Kalpana Surendranath[4], Daniela Hühn[1], Rahul Bhowmick[5], Victoria Marini[2], Stefano Ferrari[1], Ian D. Hickson[5], Lumir Krejci[2,3,6,*] and Pavel Janscak[1,*,¶]

[1]Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland; [2]Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5/A7, Brno 62500, Czech Republic; [3]International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, Brno 656 91, Czech Republic; [4]Department of Biomedical Sciences, University of Westminster, 115 New Cavendish Street, London W1W 6UW, United Kingdom; [5]Center for Chromosome Stability and Center for Healthy Aging, Department of Cellular and Molecular Medicine, University of Copenhagen, Panum Insitute Building 18.1, Blegdamsvej 3B, 2200 Copenhagen N, Denmark; [6]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5/A4, 625 00, Brno, Czech Republic

*Correspondence: pjanscak@imcr.uzh.ch and lkrejci@chemi.muni.cz
¶ Lead Contact
[+]Co-first authors

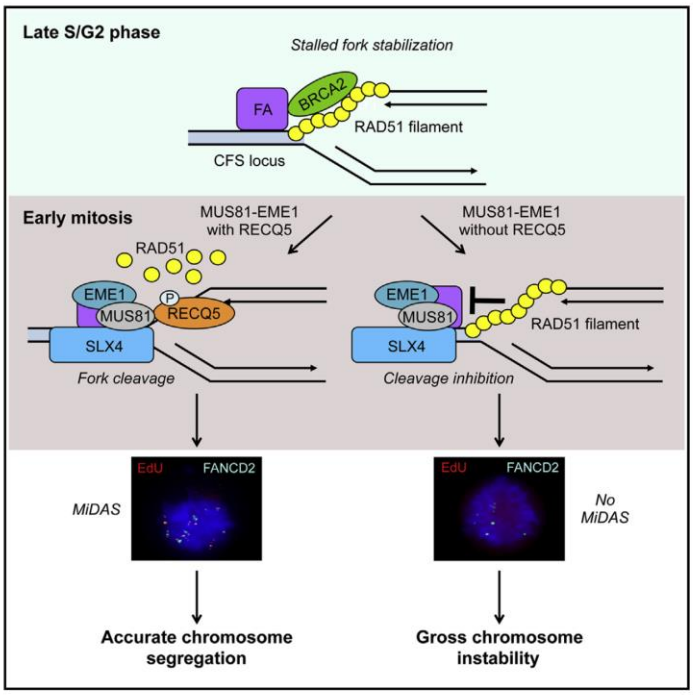
I participated in this study by performing some experiments for the analysis of ultra-fine bridge formation and contibuted to the generation of U2OS T-REx stable cell line expressing RECQ5Δ515-568.

Article

Molecular Cell

RECQ5 Helicase Cooperates with MUS81 Endonuclease in Processing Stalled Replication Forks at Common Fragile Sites during Mitosis

Graphical Abstract



Highlights

- RECQ5 binds to and stimulates MUS81-EME1 endonuclease
- RECQ5 promotes mitotic DNA synthesis (MiDAS) at CFSs
- RECQ5 removes RAD51 filaments protecting stalled replication forks at CFSs
- Expression of CFSs depends on RECQ5 phosphorylation by CDK1

Authors

Stefano Di Marco, Zdenka Hasanova, Radhakrishnan Kanagaraj, ..., Ian D. Hickson, Lumir Krejci, Pavel Janscak

Correspondence

pjanscak@imcr.uzh.ch (P.J.), lkrejci@chemi.muni.cz (L.K.)

In Brief

MUS81-EME1 endonuclease cleaves late replication intermediates at common fragile sites in early mitosis to trigger DNA-repair synthesis, ensuring faithful chromosome segregation. Di Marco et al. show that RECQ5 eliminates RAD51 filaments from these structures to facilitate the access of MUS81-EME1 to DNA.

Di Marco et al., 2017, Molecular Cell 66, 658–671
June 1, 2017 © 2017 Elsevier Inc.
http://dx.doi.org/10.1016/j.molcel.2017.05.006

CellPress

Molecular Cell
Article

RECQ5 Helicase Cooperates with MUS81 Endonuclease in Processing Stalled Replication Forks at Common Fragile Sites during Mitosis

Stefano Di Marco,[1,7] Zdenka Hasanova,[2,7] Radhakrishnan Kanagaraj,[1,7,8] Nagaraja Chappidi,[1,7] Veronika Altmannova,[2,3,7] Shruti Menon,[1] Hana Sedlackova,[2] Jana Langhoff,[1] Kalpana Surendranath,[4] Daniela Hühn,[1] Rahul Bhowmick,[5] Victoria Marini,[2] Stefano Ferrari,[1] Ian D. Hickson,[5] Lumir Krejci,[2,3,6,*] and Pavel Janscak[1,9,*]

[1]Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[2]Department of Biology, Faculty of Medicine, Masaryk University, Kamenice 5/A7, Brno 62500, Czech Republic
[3]International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, Brno 656 91, Czech Republic
[4]Department of Biomedical Sciences, University of Westminster, 115 New Cavendish Street, London W1W 6UW, UK
[5]Center for Chromosome Stability and Center for Healthy Aging, Department of Cellular and Molecular Medicine, University of Copenhagen, Panum Insitute Building 18.1, Blegdamsvej 3B, 2200 Copenhagen N, Denmark
[6]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5/A4, 625 00, Brno, Czech Republic
[7]These authors contributed equally
[8]Present address: The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK
[9]Lead Contact
*Correspondence: pjanscak@imcr.uzh.ch (P.J.), lkrejci@chemi.muni.cz (L.K.)
http://dx.doi.org/10.1016/j.molcel.2017.05.006

SUMMARY

The MUS81-EME1 endonuclease cleaves late replication intermediates at common fragile sites (CFSs) during early mitosis to trigger DNA-repair synthesis that ensures faithful chromosome segregation. Here, we show that these DNA transactions are promoted by RECQ5 DNA helicase in a manner dependent on its Ser727 phosphorylation by CDK1. Upon replication stress, RECQ5 associates with CFSs in early mitosis through its physical interaction with MUS81 and promotes MUS81-dependent mitotic DNA synthesis. RECQ5 depletion or mutational inactivation of its ATP-binding site, RAD51-interacting domain, or phosphorylation site causes excessive binding of RAD51 to CFS loci and impairs CFS expression. This leads to defective chromosome segregation and accumulation of CFS-associated DNA damage in G1 cells. Biochemically, RECQ5 alleviates the inhibitory effect of RAD51 on 3′-flap DNA cleavage by MUS81-EME1 through its RAD51 filament disruption activity. These data suggest that RECQ5 removes RAD51 filaments stabilizing stalled replication forks at CFSs and hence facilitates CFS cleavage by MUS81-EME1.

INTRODUCTION

The progression of replication forks is frequently impaired by various physical obstacles such as unrepaired DNA lesions, active transcription complexes, or DNA secondary structures. Moreover, replication slows down globally upon activation of oncogenes, which can deregulate the replication process during tumorigenesis (Gaillard et al., 2015; Zeman and Cimprich, 2014). Slowing or arrest of replication fork progression is commonly referred to as replication stress and has serious implications for genome stability and cell survival (Zeman and Cimprich, 2014).

Several genomic regions, referred to as fragile sites, are very sensitive to replication stress, and their instability is frequently associated with many human cancers. Among these are the so-called common fragile sites (CFSs) that are replicated during late S phase and recurrently give rise to gaps or breaks visible on condensed metaphase chromosomes upon partial inhibition of DNA synthesis, a phenomenon termed "expression" of CFSs (Durkin and Glover, 2007). The instability of CFSs arises from the fact that these genomic loci are difficult to replicate due to a regional paucity of replication initiation events (Debatisse et al., 2012). Under-replicated DNA regions at CFSs hamper proper chromosome segregation, ultimately causing DNA breakage, which induces DNA damage response in newly born G1 cells manifested by the formation of CFS-associated 53BP1 nuclear bodies (Lukas et al., 2011) and may lead to gross chromosomal rearrangements or loss of heterozygosity (Sofueva et al., 2011).

Recent studies have shown that CFS expression is an active process that depends on the structure-specific endonuclease MUS81-EME1 and promotes the faithful segregation of sister chromatids (Naim et al., 2013; Ying et al., 2013). The MUS81-EME1 heterodimer cleaves branched DNA structures with an exposed 5′ end at the junction, such as 3′-flap structures, replication forks, and nicked Holliday junctions (HJs) (Ciccia et al., 2003; Wyatt et al., 2013). In cells, it exists as a part of a large protein complex that contains other structure-specific nucleases, namely XPF-ERCC1 and SLX1-SLX4, where SLX4 serves as a scaffold (Svendsen et al., 2009). MUS81-EME1 is recruited to CFSs in a SLX4-dependent manner upon entry into prophase

(Minocherhomji et al., 2015). This correlates with the enhanced formation of the SLX1-SLX4-MUS81-EME1 complex in early mitosis driven by CDK1 activity (Wyatt et al., 2013). During prophase, the nuclease activity of MUS81-EME1 promotes POLD3-dependent DNA synthesis at CFSs, which serves to minimize chromosome mis-segregation and non-disjunction under conditions of replication stress (Minocherhomji et al., 2015). This mitotic DNA synthesis, known as MiDAS, likely occurs via a RAD52-mediated break-induced replication process (Bhowmick et al., 2016). MiDAS is detectable at the majority of CFS-originating breaks or gaps on metaphase chromosomes, suggesting that these cytogenetically defined loci comprise DNA synthesized at early mitosis rather than DNA strand breaks (Minocherhomji et al., 2015).

The tumor suppressors BRCA1 and BRCA2, in cooperation with the Fanconi anemia (FA) pathway, stabilize RAD51 filaments on single-stranded DNA (ssDNA) regions at stalled replication forks to protect the nascent DNA strands from MRE11-dependent degradation (Schlacher et al., 2011, 2012). However, if these stabilized forks persist until mitosis, as is the case for late replication intermediates at CFSs, RAD51 filaments would impair their processing by MUS81-EME1 and hence interfere with chromosome segregation. In fact, RAD51 over-expression promotes aneuploidy and chromosomal rearrangements in mouse embryonic stem cells (Richardson et al., 2004). Conversely, RAD51 depletion increases CFS expression (Chan et al., 2009).

RECQ5 (also termed RECQL5), a member of the RecQ family of DNA helicases, can dismantle RAD51 filaments in a reaction dependent on its ATPase activity and direct binding to RAD51 (Hu et al., 2007; Schwendener et al., 2010). In addition to its role in the regulation of RAD51 filament assembly during homologous recombination (Paliwal et al., 2014), RECQ5 has been implicated in the maintenance of CFS stability (Saponaro et al., 2014). Here, we provide several lines of evidence to suggest that RECQ5 eliminates RAD51 filaments from stalled replication forks at CFSs to facilitate their processing by MUS81-EME1 in early mitosis.

RESULTS

RECQ5 Binds to MUS81 and Stimulates Its DNA Cleavage Activity

We have shown previously that the yeast Srs2 helicase directly stimulates the endonuclease activity of Mus81-Mms4 (Chavdarova et al., 2015). Therefore, we sought to examine whether the human MUS81-EME1 endonuclease interacts functionally with the RECQ5 helicase, which is a potential ortholog of Srs2 in mammals (Hu et al., 2007). We found that RECQ5, in the absence of ATP, significantly enhanced 3′-flap DNA cleavage by MUS81-EME1 (Figures 1A–1D). RECQ5 also markedly stimulated cleavage of other MUS81-EME1 substrates including double-stranded fork and nicked HJ, whereas it did not stimulate cleavage of the Y-structure, intact HJ, or 5′-flap structures (Figures 1C and 1D), which are known to be poor substrates for MUS81-EME1 (Ciccia et al., 2003).

To test whether RECQ5 and MUS81-EME1 interact physically, as is seen with their yeast counterparts (Chavdarova et al., 2015), recombinant forms of these proteins were subjected to glutathione S-transferase (GST) pull-down assays. We found that RECQ5 was bound to GST-tagged MUS81-EME1 but not to GST alone (Figure 1E). Moreover, RECQ5 tightly associated with GST-tagged MUS81 alone, suggesting that its interaction with the MUS81-EME1 heterodimer is mediated through the MUS81 subunit (Figure S1A). Consistent with this finding, we observed that MUS81 specifically co-immunoprecipitated with RECQ5 from extracts of human U2OS cells (Figure 1F).

To map the MUS81-interaction domain in RECQ5, we generated a series of N- and C-terminally truncated variants of RECQ5 and tested them for the ability to interact with MUS81. The results suggested that the formation of a stable complex between RECQ5 and MUS81 depends on at least two domains of RECQ5. These are located within the N-terminal helicase domain and the region between the helicase domain and the RAD51-binding domain (Figures S1B–S1D). To further map the region of the MUS81 interaction within RECQ5, we constructed and analyzed a series of internal deletions of RECQ5 between its helicase and RAD51-binding domains. This identified a region between amino acids 515–570 as being important for the interaction (Figures 1G and 1H; Figure S1B). This region overlaps with the so-called KIX domain (amino acids 545–611) that mediates the interaction of RECQ5 with the jaw domain of the catalytic subunit of RNA polymerase II (RNAPII) (Islam et al., 2010; Kassube et al., 2013). We also found that RECQ5Δ515-568 did not bind to MUS81 in vitro and failed to stimulate the endonuclease activity of MUS81-EME1 on 3′-flap DNA substrate (Figure 1I; Figure S1E). On the other hand, RECQ5Δ515-568 was still able to form a complex with the hyper-phosphorylated (elongating) form of RNAPII (RNAPIIO; Figure S1F).

Together, these data indicate that RECQ5 stimulates the catalytic activity of the MUS81 endonuclease through a direct physical interaction and point to a possible role for RECQ5 as a MUS81-EME1 co-factor in vivo.

RECQ5 Promotes Expression of Common Fragile Sites

To explore the functional relationship between RECQ5 and MUS81 in human cells, we investigated the role of RECQ5 in MUS81-mediated processing of stalled replication forks at CFSs. To this end, we evaluated the effect of RECQ5 depletion on the appearance of the characteristic chromatid breaks and gaps on prometaphase chromosome spreads from cells exposed to a low dose of the DNA replication inhibitor aphidicolin (APH). Treatment of mock-depleted U2OS cells with APH stimulated expression of chromosome fragility, with about two chromatid breaks or gaps per chromosome spread (Figures 2A–2C). Upon MUS81 depletion, the number of these breaks and gaps decreased by half (Figures 2A–2C). Importantly, depletion of RECQ5 or both RECQ5 and MUS81 suppressed the formation of APH-induced chromatid breaks and gaps to the same extent as MUS81 depletion (Figures 2A–2C), implying that RECQ5 is involved in MUS81-mediated processing of CFSs.

Next, by measuring incorporation of 5-ethynyl-2′-deoxyuridine (EdU) on metaphase chromosome spreads, we evaluated the effect of RECQ5 depletion on APH-induced MiDAS that occurs predominantly at CFSs upon their cleavage by MUS81-EME1 (Minocherhomji et al., 2015). In agreement with previously

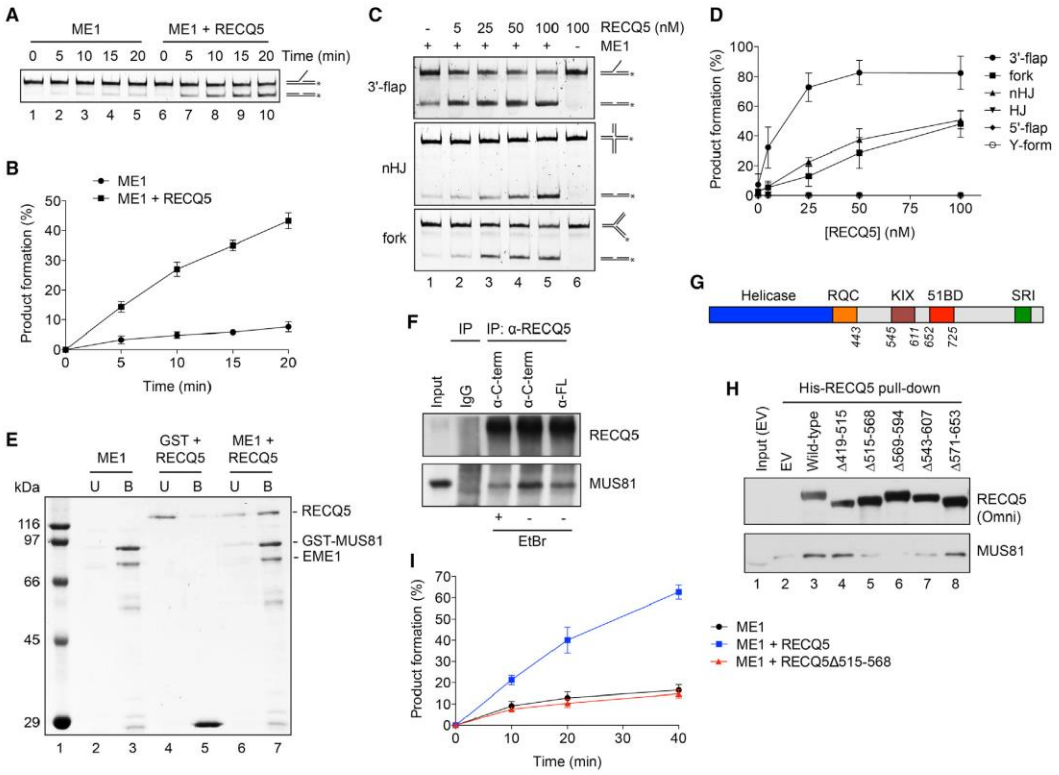


Figure 1. Physical and Functional Interaction between MUS81-EME1 Endonuclease and RECQ5

(A) Time course of cleavage of 3′-flap DNA substrate (6 nM) with MUS81-EME1 (ME1; 0.2 nM) in the presence or absence of RECQ5 (5 nM).

(B) Quantification of (A).

(C) Effect of increasing concentrations of RECQ5 on cleavage of indicated DNA substrates (6 nM) by MUS81-EME1 (0.2 nM). Reactions were incubated at 37°C for 20 min.

(D) Quantification of (C).

(E) Physical interaction of RECQ5 with MUS81-EME1. RECQ5 was incubated for 30 min with GST-tagged MUS81-EME1 or GST alone pre-bound to glutathione Sepharose beads. Fractions of unbound (U) and bound (B) proteins were resolved by SDS-PAGE.

(F) Complex formation between MUS81 and RECQ5 in human cells. Immunoprecipitation (IP) of RECQ5 from a total extract of U2OS cells was performed using rabbit polyclonal antibodies against a C-terminal (α-C-term) region of RECQ5 (675-991) and the full-length (α-FL) RECQ5, respectively. Where indicated, extracts were supplemented with ethidium bromide (EtBr; 50 μg/ml) to disrupt DNA-protein interactions.

(G) Schematic representation of RECQ5 domain organization. RQC, RecQ C-terminal domain; KIX, kinase-inducible domain interacting; 51BD, RAD51-binding domain; SRI, Set2-Rpb1 interaction motif. Numbers indicate boundaries of the individual domains.

(H) Mapping the MUS81-interaction domain of RECQ5. Indicated RECQ5 variants were expressed ectopically in HEK293 cells as N-terminal fusions with a 6xHis-Xpress epitope tag. Cell extracts were incubated with Ni-NTA beads, and bound proteins were analyzed by western blotting.

(I) Effect of wild-type RECQ5 and RECQ5Δ515-568 (25 nM) on cleavage of 3′-flap DNA substrate (4 nM) by MUS81-EME1 (0.2 nM). Reactions were carried out as in (A). For (B), (D), and (I), data are means of at least three independent experiments. Error bars show standard deviation (SD).

See also Figure S1.

published data, metaphase spreads of mock-depleted U2OS cells showed numerous EdU foci localizing to DAPI-negative breaks or gaps in condensed chromosomes (Figure 2D). MUS81 depletion caused a marked reduction in the frequency of EdU foci on metaphase spreads (Figure 2E; Figure S2A). Importantly, RECQ5 depletion or co-depletion with MUS81 inhibited EdU incorporation to the same extent as depletion of MUS81 alone (Figure 2E; Figure S2A). We also tested the effect of RECQ5 depletion on the frequency of FANCD2 twin foci (a marker of stalled forks at CFSs) co-localizing with EdU foci in prometaphase cells (Minocherhomji et al., 2015). We observed that RECQ5-depleted cells exhibited a significant reduction in the number of EdU-positive FANCD2 foci per nucleus as compared to mock-depleted cells (Figures S2B and S2C). Taken

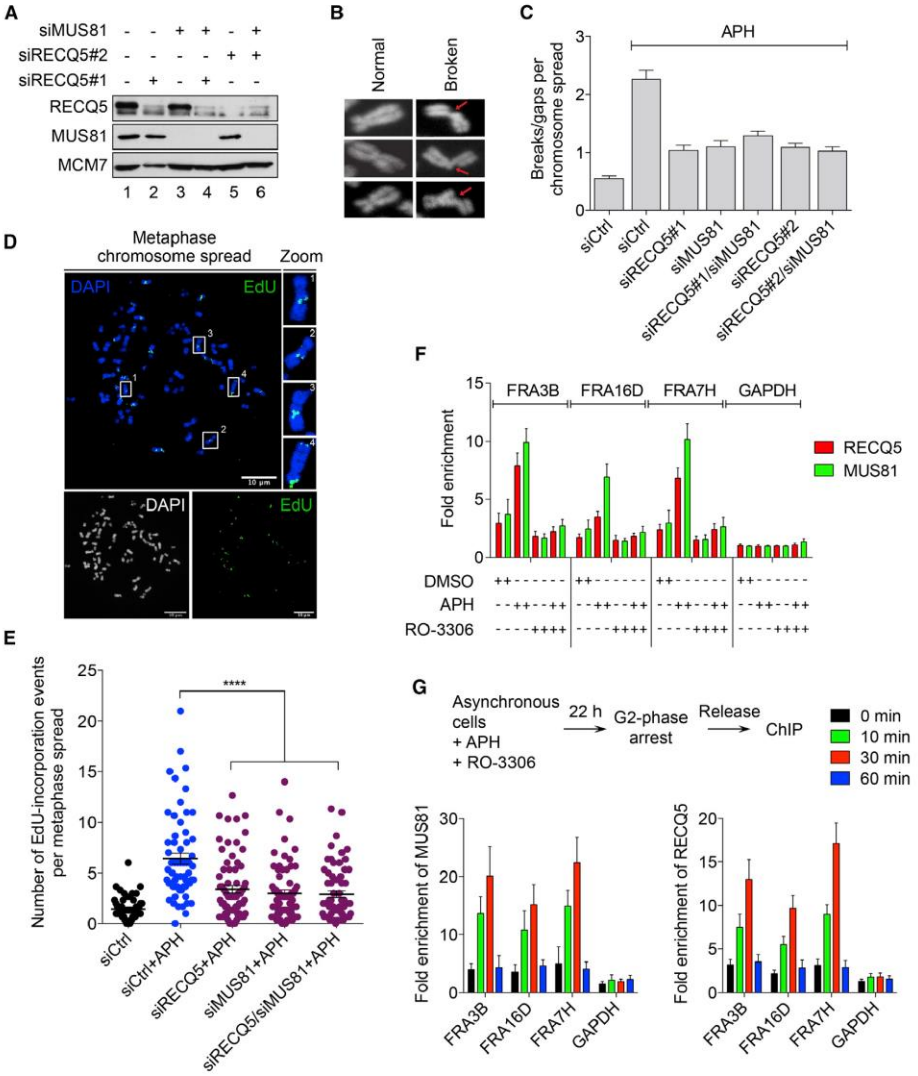


Figure 2. RECQ5 Associates with CFSs and Is Required for CFS Expression

(A) Western blot analysis of extracts from U2OS cells transfected with indicated siRNAs. Lane 1, control siRNA (siCtrl).

(B) Examples of intact and broken prometaphase chromosomes of U2OS cells treated with aphidicolin (APH; 0.2 μM) for 24 hr. Nocodazole (200 ng/ml) was added 5 hr before collection. Arrows denote chromatid breaks.

(C) Quantification of CFS expression in U2OS cells transfected with indicated siRNAs. Data are means of three independent experiments. Error bars show SEM.

(D) EdU incorporation on metaphase chromosomes. U2OS cells were treated with APH (0.4 μM) and RO-3306 (9 μM) for 16 hr (late G2 arrest) and then released into fresh medium containing EdU (20 μM) and Colcemid (0.1 μg/ml) for a further 60 min. Selected regions are magnified in numbered panels.

(E) Quantification of EdU incorporation on metaphase chromosomes for mock-, MUS81-, RECQ5-, and MUS81/RECQ5-depleted cells treated as in (D). For each condition, at least 150 metaphases were analyzed in three different experiments. The graph is scattered dot plot with black lines at mean. Error bars show SEM. ****$p < 0.0001$ (Mann-Whitney test).

(F) Specific binding of RECQ5 and MUS81 to CFSs in response to replication stress. U2OS cells were treated with APH (0.4 μM) or DMSO for 24 hr. Where indicated, RO-3306 (9 μM) was added together with APH/DMSO. Cross-linked chromatin was subjected to chromatin immunoprecipitation (ChIP) with RECQ5

(legend continued on next page)

together, these data suggest that RECQ5 promotes DNA synthesis at CFSs in mitosis.

RECQ5 Associates with CFSs during Early Mitosis upon Replication Stress

To investigate whether RECQ5 has a direct role in promoting CFS expression, we analyzed whether it physically associates with CFS loci in a manner dependent on replication stress. By chromatin immunoprecipitation (ChIP), we found that both MUS81 and RECQ5 were significantly enriched at three different CFSs tested (FRA3B, FRA16D, and FRA7H), as compared to the GAPDH locus (Figure 2F). Notably, the binding of MUS81 and RECQ5 to CFSs, but not to the GAPDH locus, was increased by 2- to 3-fold upon exposure of cells to a low dose of APH (Figure 2F), suggesting that MUS81 and RECQ5 associate with stalled replication forks at CFSs. In cells arrested at G2/M transition by CDK1 inhibition with RO-3306, the binding of RECQ5 and MUS81 to CFSs was abolished (Figure 2F), suggesting that these proteins are recruited to CFSs upon entry into mitosis. To confirm this, we analyzed the binding of RECQ5 and MUS81 to CFSs in APH-treated cells upon their release from G2/M arrest induced by RO-3306. We observed that binding of both proteins to CFSs gradually increased within 30 min following release, whereas it was barely detectable at 60 min following release (Figure 2G). Taken together, these data suggest that RECQ5 and MUS81 are recruited to CFSs in early mitosis in cells exposed to replication stress and dissociate from these loci in late mitosis.

RECQ5 Suppresses CFS-Specific Sister Chromatid Non-disjunction

To investigate whether RECQ5 depletion phenocopies MUS81 deficiency with regard to defective chromosome segregation (Naim et al., 2013; Ying et al., 2013), we first measured the frequency of DAPI-stained anaphase bridges and micronuclei in U2OS cells depleted of RECQ5, MUS81, or both. The frequency of these cytological markers of chromosome mis-segregation was significantly increased not only in MUS81-depleted cells but also in RECQ5-depleted cells, as compared with mock-treated cells (Figures 3A–3C). Simultaneous knockdown of RECQ5 and MUS81 had no additive effect in comparison with knockdown of either protein alone (Figures 3B and 3C). Second, we evaluated the effect of RECQ5 depletion on the frequency of APH-induced ultrafine bridges (UFBs) visualized by immunofluorescence staining of PICH (Chan et al., 2009). It has been shown that MUS81 depletion specifically increases UFBs positive for FANCD2 foci that mark CFS loci (Ying et al., 2013). Similarly, we found that the frequency of FANCD2-positive UFBs increased 2-fold in RECQ5-depleted cells as compared to mock-depleted cells (Figures 3D and 3E). Like MUS81 depletion (Ying et al., 2013), RECQ5 depletion did not affect the frequency of FANCD2-negative UFBs that originate from centromeric regions (Figure 3E). Simultaneous depletion of RECQ5 and MUS81 did not lead to a further increase in the frequency of FANCD2-positive UFBs as compared to single depletions (Figure 3E), supporting the notion that RECQ5 and MUS81 act in the same pathway for promoting sister chromatid disjunction.

The CFS-specific sister chromatid non-disjunction caused by MUS81 deficiency results in a marked increase in the number of 53BP1 nuclear bodies associated with CFS sequences in newly formed G1 daughter cells (Naim et al., 2013; Ying et al., 2013). We therefore tested the effect of RECQ5 depletion on the frequency of 53BP1 bodies in U2OS cells. We observed that RECQ5 depletion significantly increased the percentage of G1 cells exhibiting more than three 53BP1 bodies (Figures 3F and 3G). Depletion of MUS81 or co-depletion of RECQ5 and MUS81 brought about the same phenotype as depletion of RECQ5 (Figures 3F and 3G). To prove that RECQ5 depletion caused formation of 53BP1 nuclear bodies at CFS loci, we performed ChIP assay for 53BP1 with U2OS cells exposed to low dose of APH for 24 hr. As expected, we observed that APH treatment resulted in 53BP1 enrichment on CFSs but not on the GAPDH gene (Figure 3H). Depletion of RECQ5 or MUS81 caused a marked increase in 53BP1 binding to CFSs both in mock- and APH-treated cells (Figure 3H). Importantly, simultaneous knockdown of RECQ5 and MUS81 had no additive effect in comparison with knockdown of either protein (Figure 3H).

RECQ5 Recruitment to CFSs Is Dependent on Its Interaction with MUS81

Identification of the RECQ5 interaction with MUS81 prompted us to test the requirement of MUS81 for the recruitment of RECQ5 to CFSs. We observed that binding of RECQ5 to CFSs was impaired in MUS81-depleted cells both in the absence and the presence of APH (Figure 4A). On the contrary, depletion of RECQ5 caused a mild increase in MUS81 binding to CFSs (Figure 4B). These data suggest that MUS81 is required for the recruitment of RECQ5 to CFSs. To prove that RECQ5 recruitment to CFSs is mediated through its physical interaction with MUS81, U2OS cells were transiently transfected with vectors expressing either GFP-tagged wild-type RECQ5 or the MUS81-binding-deficient mutant of RECQ5 (RECQ5Δ515-568) (Figure 4C). ChIP using an anti-GFP antibody showed that, unlike wild-type RECQ5, RECQ5Δ515-568 was not enriched on CFSs upon APH-induced replication stress (Figure 4D), indicating that the recruitment of RECQ5 to CFSs in mitosis is dependent on RECQ5-MUS81 interaction.

To investigate whether the observed physical interaction between RECQ5 and MUS81 was required for CFS expression, we established stable U2OS cell lines expressing FLAG-tagged and siRNA-resistant forms of either wild-type RECQ5 or RECQ5Δ515-568 from a doxycycline-regulatable CMV promoter. By transfecting these cells with either control siRNA

and MUS81 antibodies, respectively, or control IgG followed by quantitative real-time PCR analysis using primers for three different CFSs (FRA3B, 16D, and 7H) and one control locus (GAPDH).

(G) Specific binding of RECQ5 and MUS81 to CFSs during early mitosis under conditions of replication stress. U2OS cells were treated with APH and RO-3306 for 22 hr and then released into mitosis for different periods of time as indicated. ChIP assay was performed as in (F). For (F) and (G), data are means of three independent experiments. Error bars show SD.

See also Figure S2.

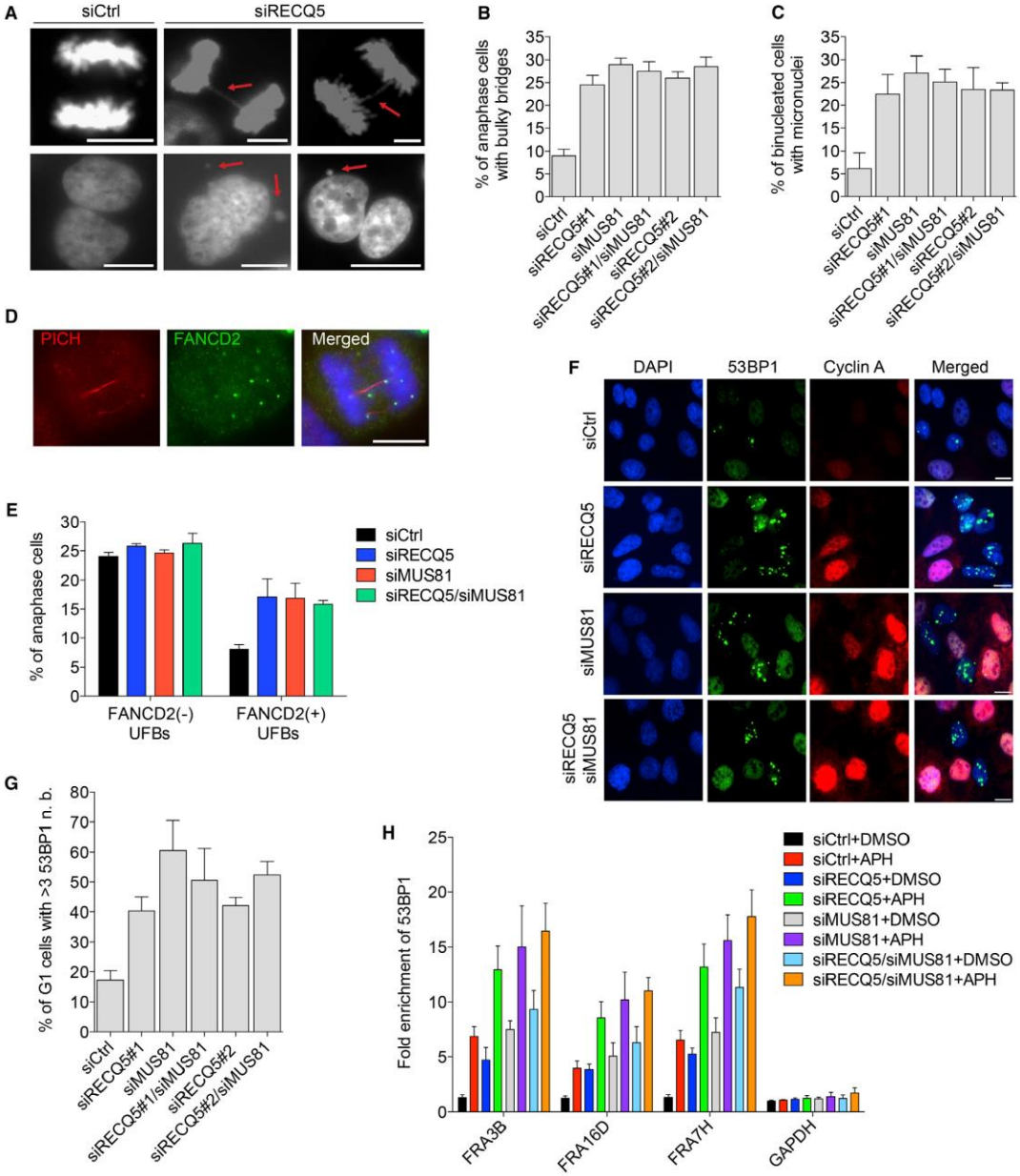


Figure 3. RECQ5 Promotes Faithful Chromosome Segregation

(A) Examples of bulky, DAPI-positive anaphase bridges (top row) and micronuclei (bottom row) scored in RECQ5-depleted U2OS cells. Red arrows indicate DAPI-positive bridges and micronuclei.

(B and C) Quantification of the frequency of DAPI-positive anaphase bridges (B) and micronuclei (C) in U2OS cells transfected with indicated siRNAs. Data are means of three independent experiments. Error bars show SD.

(legend continued on next page)

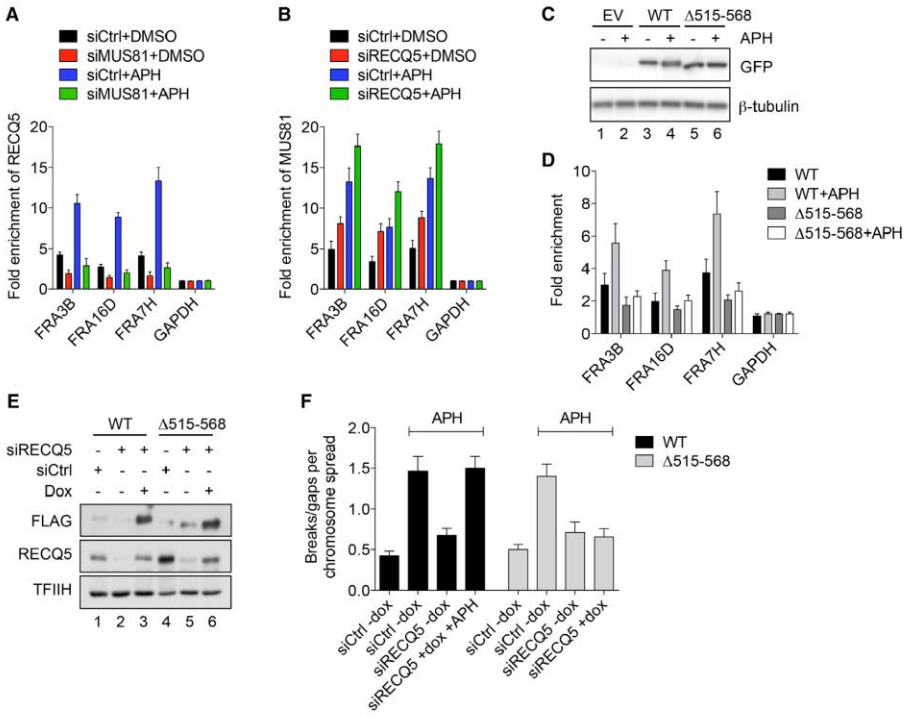


Figure 4. CFS Expression Depends on Physical Interaction between RECQ5 and MUS81

(A) Effect of MUS81 depletion on RECQ5 binding to CFSs.

(B) Effect of RECQ5 depletion on MUS81 binding to CFSs. For (A) and (B), ChIP assays were performed as in Figure 2F. Data are means of three independent experiments. Error bars show SD.

(C) Western blot analysis of extracts of U2OS cells transiently transfected with plasmids expressing GFP-tagged wild-type (WT) RECQ5 or RECQ5Δ515-568. EV, empty vector.

(D) U2OS cells transfected with plasmids as in (C) were treated with APH (0.4 μM) or DMSO for 24 hr followed by ChIP using anti-GFP antibody. Fold enrichment was calculated relative to mock-transfected cells (empty vector). Data are means of three independent experiments. Error bars show SD.

(E) Western blot analysis of extracts of U2OS T-REx cells harboring WT or Δ515-568 forms of siRNA-resistant RECQ5-FLAG fusion gene controlled by a doxycycline-inducible promoter. Following transfection (24 h) of indicated siRNAs, doxycycline (Dox) or DMSO was added with fresh medium for a further 48 hr.

(F) Quantification of CFS expression for cells in (E) cultured under indicated conditions. Cells were treated with APH (0.2 μM) for 24 hr. Nocodazole (200 ng/ml) was added 5 hr before collection. Data are means of three independent experiments. Error bars show SEM.

(siCtrl) or RECQ5 siRNA (siRECQ5) in the absence or presence of doxycycline (Dox), we could generate the following conditions: (1) expression of endogenous RECQ5 (siCtrl/-dox), (2) RECQ5 depletion (siRECQ5/-dox), and (3) substitution of endogenous RECQ5 with RECQ5 variant of interest (siRECQ5/+dox) (Figure 4E). The results showed that replacement of endogenous RECQ5 with RECQ5Δ515-568, but not wild-type RECQ5, impaired CFS expression (Figure 4F), thus providing further

(D) Example of CFS-associated ultrafine bridges (UFBs) scored in RECQ5-depleted U2OS cells treated with 0.4 μM APH for 24 hr. CFS-associated UFBs are marked by PICH on the bridge (red) and FANCD2 foci (green) on the bridge termini.

(E) Quantification of the frequency of centromeric (FANCD2-negative) and CFS-associated (FANCD2-positive) UFBs in U2OS cells transfected with indicated siRNA. Data are means of three independent experiments. Error bars show SD.

(F) Examples of G1 phase cells (cyclin A-negative, red) containing 53BP1 nuclear bodies (green).

(G) Quantification of G1-phase specific 53BP1 nuclear bodies in U2OS cells transfected with indicated siRNAs. Percentage of G1 cells with >3 nuclear bodies (n.b.) is plotted. Data are means of four independent experiments. Error bars show SD. At least 600 cyclin A-negative cells were analyzed for each condition.

(H) 53BP1 bodies in RECQ5-depleted U2OS cells are associated with CFS loci. Cells transfected with indicated siRNAs were treated with APH (0.4 μM) or DMSO for 24 hr. Cross-linked chromatin was subjected to ChIP assay using 53BP1 antibody or control IgG. Data are means of three independent experiments. Error bars show SD.

evidence for a direct involvement of RECQ5 in MUS81-mediated processing of late replication intermediates in mitosis.

RECQ5 Eliminates RAD51 from Stalled Replication Forks at CFSs to Facilitate Their Processing by MUS81-EME1

As RECQ5 is able to disrupt RAD51 nucleoprotein filaments (Hu et al., 2007), we reasoned that it might remove RAD51 from stalled replication forks at CFSs, allowing DNA cleavage by MUS81-EME1. Thus, we analyzed the effect of RECQ5 depletion on RAD51 occupancy at CFSs in U2OS cells exposed to replication stress generated by a low dose of APH. ChIP assay showed that, in mock-depleted cells, APH treatment enhanced binding of RAD51 to CFSs but not to the GAPDH locus, consistent with the assumption that RAD51 protects stalled forks at CFSs (Figure 5A). Importantly, depletion of RECQ5 increased RAD51 binding specifically to CFSs in both non-treated and APH-treated cells (Figure 5A). These results imply that RECQ5 regulates RAD51 occupancy on stalled replication forks at CFS loci.

Since disruption of RAD51 filaments by RECQ5 is dependent on its ATPase and RAD51-binding activities (Hu et al., 2007; Schwendener et al., 2010), we investigated the role of these activities in the control of RAD51 binding to CFSs. For this, we used stable U2OS T-REx cell lines harboring vectors for doxycycline-regulated expression of siRNA-resistant versions of RECQ5 mutants that cannot disrupt RAD51 filaments: RECQ5K58R (defective in ATP hydrolysis) and RECQ5F666A (defective in RAD51 binding). We found that replacement of endogenous RECQ5 with either of these mutants increased the binding of RAD51 to CFSs to levels detected in RECQ5-depleted cells (Figures 5B and 5C). On the contrary, ectopic expression of wild-type RECQ5 suppressed the excessive binding of RAD51 to CFSs in cells depleted for endogenous RECQ5 (Figures 5B and 5C). Substitution of endogenous RECQ5 with either of the two mutants, but not wild-type RECQ5, also compromised expression of chromosome fragility in APH-treated cells and increased the frequency of anaphase bridges and micronuclei to levels comparable with those observed in RECQ5-depleted cells (Figures 5D–5F). These results provide further evidence that RECQ5 eliminates RAD51 filaments from stalled replication forks at CFSs.

A previous study suggested that RECQ5 protects CFSs by regulating RNAPII-mediated transcription to prevent transcription stress (Saponaro et al., 2014). It has also been demonstrated that RECQ5 inhibits transcription through its KIX domain, which binds to the jaw domain of the catalytic subunit of RNAPII (Kassube et al., 2013). Therefore, we also established a stable U2OS T-REx cell line for inducible expression of a RECQ5 mutant carrying a E584D substitution in the KIX domain, which impairs its binding to RNAPII (Islam et al., 2010) (Figure S3A). By ChIP for 53BP1 binding, we found that replacement of endogenous RECQ5 with the KIX mutant did not induce DNA damage at the CFSs tested (Figures S3B and S3C). In contrast, cells expressing RECQ5 mutants that cannot disrupt RECQ5 filaments (K58R and F666A) displayed an increase in binding of 53BP1 to CFSs similar to levels detected in RECQ5-depleted cells (Figure 5G). We also observed that replacement of endogenous RECQ5 with the KIX mutant had no effect on expression of CFSs and the frequency of anaphase bridges (Figures S3D and S3E). It should be noted that the observed residual binding of RECQ5E584D to RNAPIIO (Figure S3A) is mediated by the Set2-Rpb1-interacting (SRI) domain of RECQ5 that binds to the C-terminal repeat domain of the RNAPII catalytic subunit during transcription elongation (Islam et al., 2010).

Taken together, our results provide strong evidence that RECQ5 acts during early mitosis to dismantle RAD51 filaments on stalled replication forks at CFSs and hence facilitates CFS cleavage by MUS81-EME1, which ensures CFS stability. In agreement with this conclusion, RECQ5 depletion did not impair CFS expression in RAD51-deficient cells (Figure S4).

RECQ5 Alleviates the Inhibitory Effect of RAD51 on 3′-Flap DNA Cleavage by MUS81-EME1

To investigate whether RAD51 filaments inhibit fork cleavage by MUS81-EME1, a 3′-flap DNA substrate was incubated with increasing concentrations of RAD51, followed by the addition of MUS81-EME1. We used a K133R mutant of RAD51, which binds but does not hydrolyze ATP and hence forms a stable complex with ssDNA (Chi et al., 2006), mimicking the in vivo situation where RAD51 filaments are stabilized by BRCA2 via inhibition of RAD51 ATP hydrolysis (Jensen et al., 2010). We found that RAD51K133R inhibited 3′-flap DNA cleavage by MUS81-EME1 in a concentration-dependent manner (Figures S5A and S5B). The extent of inhibition correlated with the level of binding of RAD51K133R to the DNA substrate (Figure S5C), indicating that this effect stems from the formation of RAD51 filament.

To evaluate the effect of RECQ5 on 3′-flap cleavage by MUS81-EME1 in the presence of RAD51, the DNA substrate was pre-incubated with RAD51K133R at a concentration that largely blocked the DNA-cleavage reaction (Figure 5H, compare lanes 2 and 5). Reactions were then supplemented with increasing amounts of RECQ5, followed by the addition of MUS81-EME1. We found that RECQ5 partially restored 3′-flap DNA cleavage by MUS81-EME1 in a manner dependent on ATP (Figures 5H and 5I; Figures S6A and S6B). In contrast, RECQ5 mutants defective in disrupting RAD51 filaments, namely RECQ5K58R and RECQ5F666A, were impaired in their ability to suppress the inhibitory effect of RAD51 on 3′-flap DNA cleavage by MUS81-EME1 (Figures 5H and 5I). Crucially, these mutants stimulated efficiently the DNA cleavage activity of MUS81-EME1 on naked DNA substrate and displayed normal binding to MUS81 (Figures S6C and S6D). Taken together, these data indicate that RECQ5 can eliminate RAD51 filaments from a fork structure to facilitate its cleavage by MUS81-EME1.

MUS81-Mediated Processing of Stalled Replication Forks at CFSs Is Dependent on Phosphorylation of RECQ5 by CDK1

We noticed that RECQ5 from cells synchronized in pro-metaphase by nocodazole treatment displayed a reduced electrophoretic mobility compared to that of asynchronous cells (Figure 6A; Figure S7A). This band shift was sensitive to lambda phosphatase treatment, suggesting a specific phosphorylation of RECQ5 in early mitosis (Figure 6A). To explore whether RECQ5 is phosphorylated by CDK1-Cyclin B1, we performed an in vitro kinase assay. We observed that full-length RECQ5

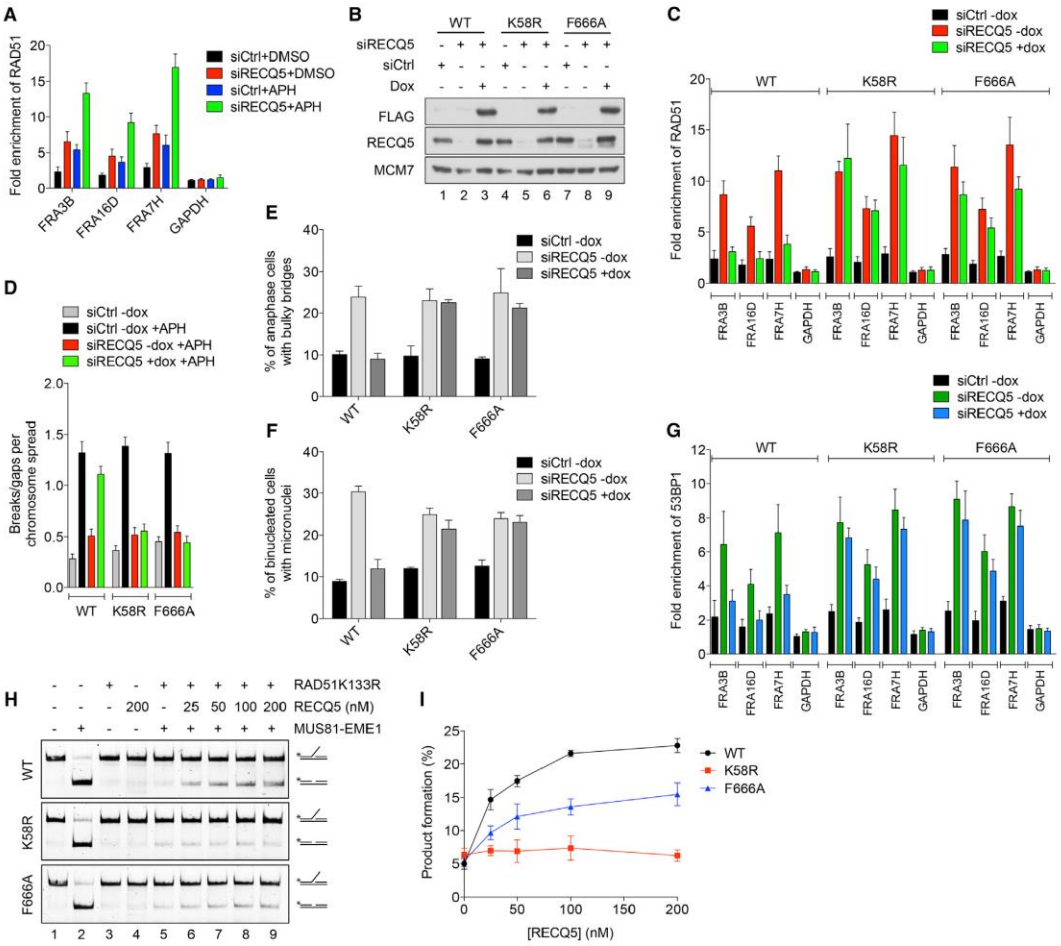


Figure 5. RAD51 Filaments Inhibit MUS81-Mediated Processing of Stalled Replication Forks at CFSs in RECQ5-Deficient Cells

(A) RECQ5 depletion causes excessive binding of RAD51 to CFSs. Cross-linked chromatin samples of mock- or RECQ5-depleted U2OS cells treated with APH (0.4 μM) or DMSO for 24 hr were subjected to ChIP assay with anti-RAD51 antibody or control IgG. Data represent means of three independent experiments. Error bars show SD.

(B) Wstern blot analysis of extracts of U2OS T-REx cells harboring wild-type (WT), K58R, or F666A forms of siRNA-resistant RECQ5-FLAG fusion gene controlled by a doxycycline-inducible promoter. Twenty-four hours after transfection of indicated siRNAs, doxycycline (Dox) or DMSO was added with fresh medium for a further 48 hr as indicated.

(C–G) Quantifications of RAD51 binding to CFSs (C), CFS expression (D), the frequency of DAPI-positive anaphase bridges (E) and micronuclei (F), and 53BP1 binding to CFSs (G) for cells in (B). Three conditions were analyzed for each cell line: siCtrl -dox (expression of endogenous RECQ5), siRECQ5 -dox (RECQ5 depletion), and siRECQ5 +dox (replacement of endogenous RECQ5 with RECQ5 variant). Data are means of three (C and G) or two (D–F) independent experiments. Error bars show SD in (C) and (E–G) and SEM in (D).

(H) RECQ5 alleviates the inhibitory effect of RAD51 filaments on 3′-flap DNA cleavage by MUS81-EME1. Fluorescently labeled 3′-flap DNA substrate (6 nM) was pre-incubated for 10 min with RAD51K133R (500 nM) in the presence of ATP (2 mM) and ATP-regenerating system. Reactions were then supplemented with indicated concentrations of wild-type (WT) or mutant (K58R, F666A) forms of RECQ5 and incubated for 10 min at 37°C, followed by addition of MUS81-EME1 (8 nM) and incubation for a further 20 min.

(I) Quantification of (H). Data are means of at least three independent experiments. Error bars show SD.

See also Figures S3, S4, S5, and S6.

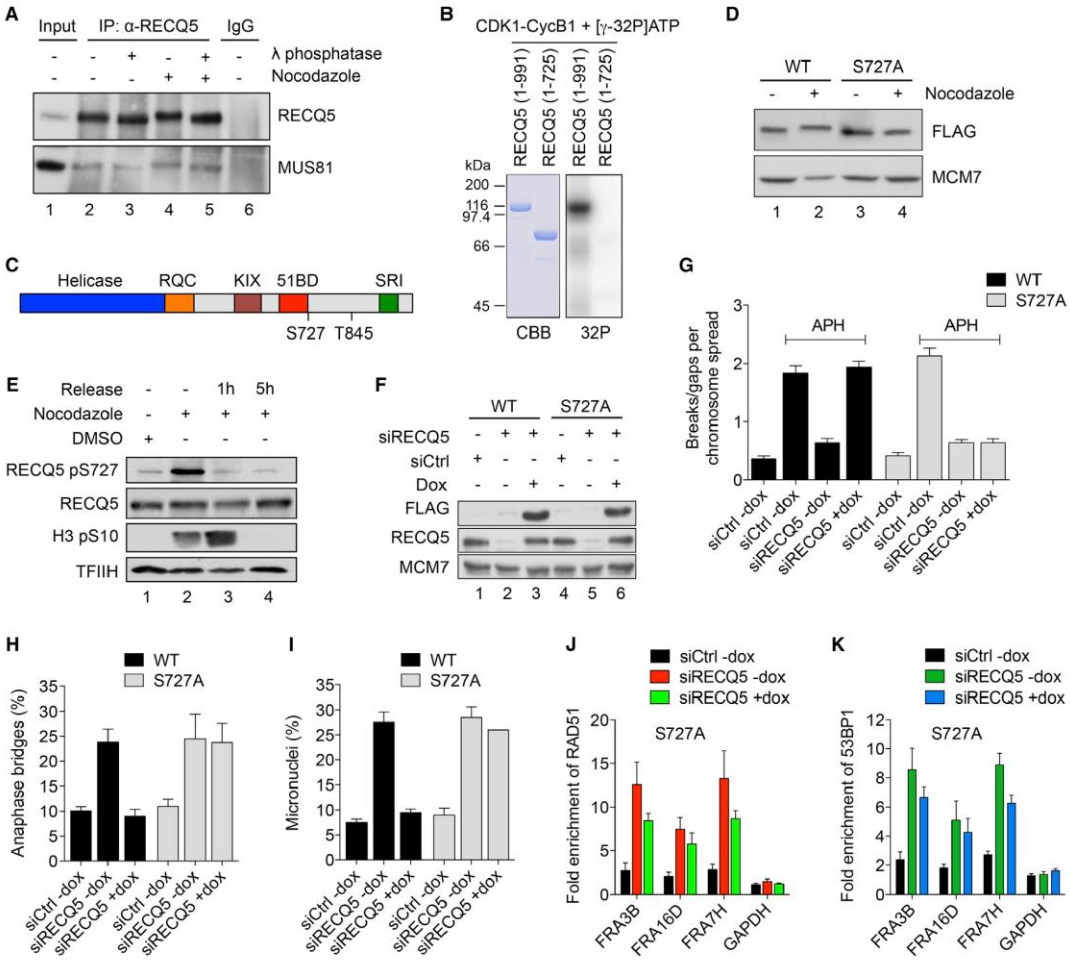


Figure 6. Dependence of CFS Expression on Phosphorylation of RECQ5 at Ser727

(A) Phosphorylation of RECQ5 in early mitosis as revealed by electrophoretic mobility shift assay. Total extracts of mock- and nocodazole-treated HeLa cells were subjected to immunoprecipitation (IP) with anti-RECQ5 antibody or control IgG followed by western blotting. Where indicated, immunoprecipitates were treated with lambda (λ) protein phosphatase.

(B) In vitro phosphorylation of RECQ5 by CDK1-cyclin B1. One microgram of purified full-length RECQ5 (1-991) or RECQ5 fragment spanning amino acids 1–725 was incubated with 2U of purified CDK1-cyclin B1 complex (CDK1/CycB1) and 0.05 mM ATP/[γ-^{32}P]ATP for 15 min at 37°C. Reactions were subjected to SDS-PAGE followed by Coomassie brilliant blue staining (CBB) and phosphor imaging (^{32}P).

(C) Location of two putative CDK consensus sites in RECQ5.

(D) Mutation of Ser727 of RECQ5 prevents its phosphorylation in early mitosis. HeLa cells were transiently transfected with expression vectors for wild-type (WT) RECQ5 or RECQ5S727A fused to a 3×FLAG tag. Cells were treated with nocodazole (+) or mock-treated (−) as indicated. Total cell extracts were analyzed by western blotting.

(E) RECQ5 is phosphorylated at Ser727 in early mitosis. U2OS cells were treated with nocodazole (200 ng/ml) for 16 hr and then released into fresh medium for 1 or 5 hr. Total cell extracts were analyzed by western blotting using antibody against the phospho-Ser727 epitope of RECQ5. Mitosis was monitored by H3 phosphorylation at Ser10.

(F) Western blot analysis of extracts of U2OS T-REx cells harboring WT or S727A forms of siRNA-resistant RECQ5-FLAG fusion gene. Twenty-four hours after siRNA transfection, doxycycline (Dox; "+") or DMSO ("−") was added with fresh medium for a further 48 hr.

(G–I) Quantification of CFS expression (G) and the frequency of DAPI-positive anaphase bridges (H) and micronuclei (I) for cells in (F) cultured under indicated conditions. Data are means of two independent experiments. Error bars show SEM in (G) and SD in (H) and (I).

(legend continued on next page)

was extensively phosphorylated by CDK1 (Figure 6B), whereas a C-terminal deletion variant of RECQ5 lacking two putative CDK consensus sites was not phosphorylated (Figures 6B and 6C). Strikingly, one of the two potential CDK sites, Ser727, which is located adjacent to the RAD51-interacting domain (Figure 6C), was found to be phosphorylated in several proteomic studies conducted with human and mouse cells (http://www.phosphosite.org/). To assess whether Ser727 is the target of phosphorylation in mitosis, we generated a non-phosphorylatable mutant of RECQ5 by replacing the serine residue with an alanine. When expressed in U2OS cells, this mutant did not show electrophoretic mobility shift upon nocodazole treatment, indicating that Ser727 is phosphorylated in early mitosis (Figure 6D). To substantiate these findings, we raised a phospho-Ser727-specific rabbit polyclonal antibody and confirmed that RECQ5 phosphorylation at Ser727 was dramatically increased upon nocodazole treatment (Figure 6E). Moreover, upon release of cells from nocodazole block, the phosphorylated form of RECQ5 was rapidly lost, suggesting that RECQ5 is actively dephosphorylated with the onset of anaphase (Figure 6E).

To explore the biological role of the RECQ5 Ser727 phosphorylation, we established a stable U2OS T-REx cell line inducibly expressing a siRNA-resistant version of FLAG-tagged RECQ5S727A (Figure 6F). We found that replacement of endogenous RECQ5 with the non-phosphorylatable mutant of RECQ5 largely abolished the expression of chromosome fragility in APH-treated cells and increased the frequency of anaphase bridges and micronuclei to levels observed in cells lacking RECQ5 (Figures 6G–6I). Moreover, cells expressing the S727A mutant displayed excessive binding of RAD51 and 53BP1 to CFSs when compared to cells expressing endogenous RECQ5 (Figures 6J and 6K). These data suggest that RECQ5 is phosphorylated by CDK1 at Ser727, and this phosphorylation is required for RECQ5-mediated disruption of RAD51 filaments on stalled replication forks at CFSs and subsequent fork cleavage by MUS81-EME1.

DISCUSSION

We have provided evidence that RAD51 filaments inhibit MUS81-mediated cleavage of stalled replication forks at CFSs that triggers DNA-repair synthesis at these loci in early mitosis to ensure the accurate transmission of genetic material to daughter cells (Minocherhomji et al., 2015; Naim et al., 2013; Ying et al., 2013). Moreover, we have identified RECQ5 as a factor that suppresses CFS-associated chromosome mis-segregation and non-disjunction by counteracting the inhibitory effect of RAD51 on MUS81-mediated cleavage of CFSs. Our results demonstrate that RECQ5 is recruited to CFSs by MUS81 at the onset of mitosis and disrupts RAD51 filaments formed at these loci to facilitate the access of MUS81-EME1 to the DNA. Thus, our study reveals the molecular mechanism underlying the coordination between the FA-BRCA1/2-RAD51 fork stabilization pathway (Schlacher et al., 2011, 2012) and MUS81-mediated resolution of late replication intermediates in mitosis.

In agreement with our findings, RECQ5-deficient cells have been shown to accumulate chromosomal deletions with breakpoints at CFSs (Saponaro et al., 2014). Although this was attributed to the function of RECQ5 in regulation of RNAPII movement across genes to prevent it from stalling or arrest (Saponaro et al., 2014), it should be noted that our study has shown that ablation of the binding of RECQ5′s KIX domain to RNAPII, which is implicated in regulation of RNAPII transcription (Kassube et al., 2013), did not cause accumulation of DNA damage at CFSs, a phenotype observed in RECQ5-depleted cells. Thus, we propose that the accumulation of CFS-associated chromosomal rearrangements in RECQ5-deficient cells is, at least in part, the consequence of a failure of MUS81-mediated processing of late replication intermediates at these loci in mitosis. The involvement of RECQ5 in this process is also supported by observed synthetic lethality between loss of RECQ5 and Werner syndrome protein (Popuri et al., 2013), another RecQ DNA helicase that is known to suppress expression of CFSs possibly by preventing replication stress (Murfuni et al., 2012). Nevertheless, our data do not exclude the possibility that transcription stress caused by RECQ5 deficiency induces genomic instability at CFSs during S/G2 phases of the cell cycle or in a manner independent of DNA replication.

A number of CFSs such as FRA3B or FRA16D are located within genes that are larger than 800 kb and hence require more than one cell cycle to be transcribed (Helmrich et al., 2011). Evidence suggests that the fragility of these CFSs is driven by inevitable collisions between replication and transcription complexes, perhaps because of the formation of R-loops under conditions of replication stress (Helmrich et al., 2011). Thus, MUS81-induced DNA synthesis at CFSs could provide a mechanism for resolution of such collisions to complete transcription of the underlying genes. Of note, we have recently shown that RECQ5 promotes the resolution of collisions between transcription and replication complexes during S-phase (Urban et al., 2016). It will be interesting to see whether the resolution of these collisions proceeds via a similar mechanism as that of MUS81-mediated processing of CFSs in early mitosis. Strikingly, we have found that the helicase activity of RECQ5 counteracts the formation of RAD51 foci that form in unperturbed S-phase cells as a consequence of transcription-replication interference and likely represent unresolved replication intermediates stabilized by RAD51 filaments (Urban et al., 2016). This is reminiscent of RECQ5-mediated removal of RAD51 from CFSs, which promotes CFS cleavage by MUS81-EME1 and subsequent DNA synthesis.

Our study revealed that CFS expression and proper chromosome segregation are dependent on phosphorylation of RECQ5 at Ser727 by CDK1. The same kinase also triggers the

(J) RECQ5 Ser727 phosphorylation prevents excessive binding of RAD51 to CFSs. U2OS T-REx cells carrying the vector for expression of the S727A form of RECQ5-FLAG were treated as in (F). Samples of cross-linked chromatin were subjected to ChIP assay as in Figure 5A.
(K) Accumulation of DNA damage at CFSs in U2OS T-REx cells expressing RECQ5 S727A. Cells were cultured as in (F). The level of 53BP1 binding at CFSs was measured as in Figure 3H. For (J) and (K), data are means of three independent experiments. Error bars show SD.
See also Figure S7.

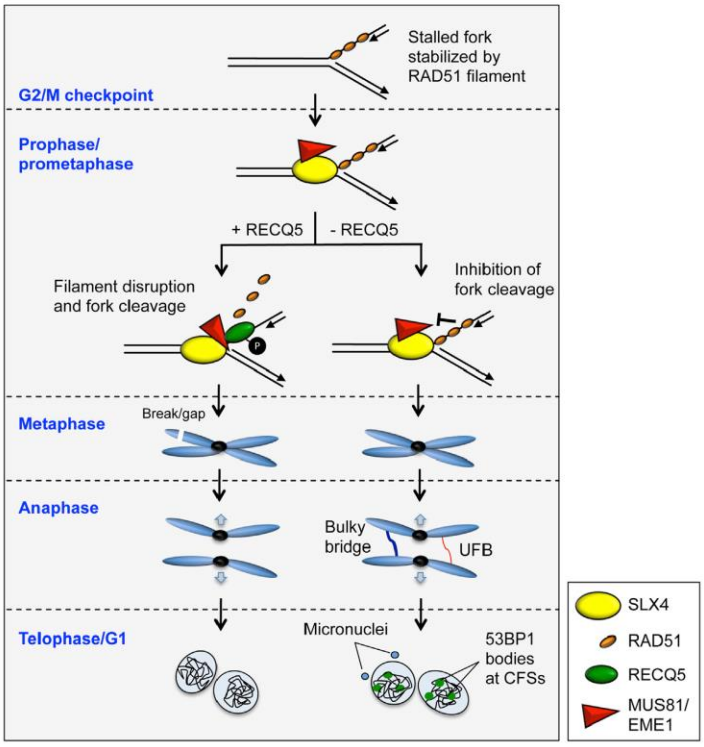


Figure 7. Model for the Role of RECQ5 in Facilitating CFS Cleavage by MUS81-EME1 Endonuclease during Early Mitosis

Perturbation of DNA replication at late-replicating common fragile sites (CFSs) leads to the assembly of RAD51 filaments on single-stranded DNA regions at the stalled forks. As such, stalled forks escape G2/M checkpoint. When cells enter mitosis, SLX4 forms complex with the MUS81-EME1 endonuclease and recruits it to CFS loci. In turn, MUS81 recruits RECQ5 through direct protein-protein interaction. Alternatively, RECQ5 is recruited in complex with MUS81-EME1. RECQ5 eliminates RAD51 filament on the leading strand of the stalled fork and stimulates fork cleavage by MUS81-EME1, which triggers DNA synthesis (not shown) required for correct sister chromatid segregation. RECQ5-mediated disruption of RAD51 filaments at CFSs is driven by phosphorylation of RECQ5 Ser727 by CDK1. RECQ5 deficiency leads to RAD51 persistency on replication intermediates, which consequently impedes chromosome segregation. The cell will inevitably go through an aberrant anaphase separation, leading to genomic instability.

cell-cycle stage-specific association of MUS81-EME1 with SLX1-SLX4, which is required for the recruitment of MUS81-EME1 to CFSs (Minocherhomji et al., 2015; Wyatt et al., 2013). Although RECQ5 recruitment to CFSs is dependent on its physical interaction with MUS81, it is not likely that this process requires RECQ5 phosphorylation because complex formation between RECQ5 and MUS81 was not enhanced in cells arrested in early mitosis by nocodazole treatment (Figure 6A). Thus, RECQ5 phosphorylation might rather play a role in the action of RECQ5 at CFSs. Interestingly, Ser727 is located in close proximity to the RAD51-binding domain of RECQ5 (Figure 6C), suggesting that its phosphorylation might contribute to the efficient disruption of RAD51 filaments by RECQ5. In agreement with this possibility, we found that RECQ5 Ser727 phosphorylation is required for the suppression of RAD51 binding to CFSs. However, results of our biochemical experiments with phosphomimetic mutant of RECQ5 (S727D) suggest that RECQ5 phosphorylation neither directly enhances its physical interaction with RAD51 nor facilitates RECQ5-mediated displacement of RAD51 from ssDNA (Figures S7B–S7G). Thus, further studies are needed to clarify the role of RECQ5 Ser727 phosphorylation in the regulation of RAD51 filament assembly in mitotic cells.

In conclusion, the results of our study are consistent with a model wherein MUS81-mediated disjunction of under-replicated DNA at CFSs is dependent on RECQ5 helicase that facilitates the access of the MUS81-EME1 endonuclease to the DNA by disrupting RAD51 filaments formed on the arms of stalled forks (Figure 7). In the absence of such activity, persistent RAD51 filaments inhibit CFS cleavage by MUS81-EME1, leading to the formation of CFS-associated UFBs. These can give rise to DNA damage that is transmitted to the newly born G1 cells (Figure 7). Our findings may have important implications for cancer therapy, as there is a tendency for RAD51 over-expression in tumor cells, leading to increased resistance to chemotherapeutics (Klein, 2008). Since RAD51 over-expression stabilizes RAD51 filaments (Schlacher et al., 2012), it is possible that RAD51-over-expressing cells exhibit an increased dependency on RECQ5 for their survival. Thus, targeting RECQ5 in these cancers with small molecule inhibitors might provide a therapeutic approach.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Protein expression and purification
  - Nuclease assays
  - D-loop assay
  - Gel retardation assay
  - Plasmid constructions
  - Cell culture

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2017.05.006.

AUTHOR CONTRIBUTIONS

S.D.M. and N.C. performed analyses of chromosome spreads, anaphase bridges, micronuclei, and 53BP1 nuclear bodies. Z.H. and V.A. performed protein purifications, the DNA cleavage experiments, and protein-protein interaction analysis. R.K. performed ChIP experiments with the help of K.S.; J.L. identified RECQ5 Ser727 phosphorylation. D.H. and J.L. performed immunoprecipitation assays. H.S. and V.M. purified proteins, performed some nuclease assays, and prepared DNA substrates. R.K. and S.F. performed in vitro kinase assays. S.M. performed UFB analysis together with S.D.M.; R.B. and I.D.H. analyzed mitotic DNA synthesis in RECQ5-depleted cells. P.J. and L.K. designed the study, analyzed the data, and wrote the manuscript. I.D.H. edited the manuscript.

ACKNOWLEDGMENTS

We thank M. Sebesta, M. Spirek, B. Stefanovie, and A. Sturzenegger for reagents. This work was supported by grants from the Swiss National Science Foundation (31003A_146206 and 31003A_166451), the Promedica Foundation (GHDE AUGK-DZZ 1226/M), and the Stiftung zur Krebsbekämpfung to P.J. and from the Czech Science Foundation (13-26629S and 17-17720S), the European Regional Development Fund – Project FNUSA-ICRC (no. CZ.1.05/1.1.00/02.0123), and the Research Support Programme (GAMU) (MUNI/M/1894/2014) to L.K. In addition, K.S. was supported by the University of Westminster's Faculty of Science and Technology, and I.D.H. was supported by grants from the Danish National Research Foundation (DNRF115), The Nordea Foundation, and the European Research Council (321717). R.K. and R.B. were recipients of postdoctoral fellowships from UBS AG and the Danish Medical Research Council, respectively.

Received: September 16, 2016
Revised: February 14, 2017
Accepted: May 4, 2017
Published: June 1, 2017

REFERENCES

Bergoglio, V., Boyer, A.S., Walsh, E., Naim, V., Legube, G., Lee, M.Y., Rey, L., Rosselli, F., Cazaux, C., Eckert, K.A., and Hoffmann, J.S. (2013). DNA synthesis by Pol η promotes fragile site stability by preventing under-replicated DNA in mitosis. J. Cell Biol. *201*, 395–408.

Bhowmick, R., Minocherhomji, S., and Hickson, I.D. (2016). RAD52 facilitates mitotic DNA synthesis following replication stress. Mol. Cell *64*, 1117–1126.

Burkovics, P., Dome, L., Juhasz, S., Altmannova, V., Sebesta, M., Pacesa, M., Fugger, K., Sorensen, C.S., Lee, M.Y., Haracska, L., and Krejci, L. (2016). The PCNA-associated protein PARI negatively regulates homologous recombination via the inhibition of DNA repair synthesis. Nucleic Acids Res. *44*, 3176–3189.

Chan, K.L., Palmai-Pallag, T., Ying, S., and Hickson, I.D. (2009). Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. Nat. Cell Biol. *11*, 753–760.

Chavdarova, M., Marini, V., Sisakova, A., Sedlackova, H., Vigasova, D., Brill, S.J., Lisby, M., and Krejci, L. (2015). Srs2 promotes Mus81-Mms4-mediated resolution of recombination intermediates. Nucleic Acids Res. *43*, 3626–3642.

Chi, P., Van Komen, S., Sehorn, M.G., Sigurdsson, S., and Sung, P. (2006). Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function. DNA Repair (Amst.) *5*, 381–391.

Ciccia, A., Constantinou, A., and West, S.C. (2003). Identification and characterization of the human mus81-eme1 endonuclease. J. Biol. Chem. *278*, 25172–25178.

Debatisse, M., Le Tallec, B., Letessier, A., Dutrillaux, B., and Brison, O. (2012). Common fragile sites: mechanisms of instability revisited. Trends Genet. *28*, 22–32.

Durkin, S.G., and Glover, T.W. (2007). Chromosome fragile sites. Annu. Rev. Genet. *41*, 169–192.

Gaillard, H., García-Muse, T., and Aguilera, A. (2015). Replication stress and cancer. Nat. Rev. Cancer *15*, 276–289.

Helmrich, A., Ballarino, M., and Tora, L. (2011). Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol. Cell *44*, 966–977.

Henricksen, L.A., Umbricht, C.B., and Wold, M.S. (1994). Recombinant replication protein A: expression, complex formation, and functional characterization. J. Biol. Chem. *269*, 11121–11132.

Hu, Y., Raynard, S., Sehorn, M.G., Lu, X., Bussen, W., Zheng, L., Stark, J.M., Barnes, E.L., Chi, P., Janscak, P., et al. (2007). RECQL5/Recql5 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments. Genes Dev. *21*, 3073–3084.

Islam, M.N., Fox, D., 3rd, Guo, R., Enomoto, T., and Wang, W. (2010). RecQL5 promotes genome stabilization through two parallel mechanisms—interacting with RNA polymerase II and acting as a helicase. Mol. Cell. Biol. *30*, 2460–2472.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. Nature *467*, 678–683.

Kanagaraj, R., Huehn, D., MacKellar, A., Menigatti, M., Zheng, L., Urban, V., Shevelev, I., Greenleaf, A.L., and Janscak, P. (2010). RECQ5 helicase associates with the C-terminal repeat domain of RNA polymerase II during productive elongation phase of transcription. Nucleic Acids Res. *38*, 8131–8140.

Kassube, S.A., Jinek, M., Fang, J., Tsutakawa, S., and Nogales, E. (2013). Structural mimicry in transcription regulation of human RNA polymerase II by the DNA helicase RECQL5. Nat. Struct. Mol. Biol. *20*, 892–899.

Klein, H.L. (2008). The consequences of Rad51 overexpression for normal and tumor cells. DNA Repair (Amst.) *7*, 686–693.

Lukas, C., Savic, V., Bekker-Jensen, S., Doil, C., Neumann, B., Pedersen, R.S., Grøfte, M., Chan, K.L., Hickson, I.D., Bartek, J., and Lukas, J. (2011). 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat. Cell Biol. *13*, 243–253.

Matulova, P., Marini, V., Burgess, R.C., Sisakova, A., Kwon, Y., Rothstein, R., Sung, P., and Krejci, L. (2009). Cooperativity of Mus81.Mms4 with Rad54 in the resolution of recombination and replication intermediates. J. Biol. Chem. *284*, 7733–7745.

Minocherhomji, S., Ying, S., Bjerregaard, V.A., Bursomanno, S., Aleliunaite, A., Wu, W., Mankouri, H.W., Shen, H., Liu, Y., and Hickson, I.D. (2015). Replication stress activates DNA repair synthesis in mitosis. Nature *528*, 286–290.

Murfuni, I., De Santis, A., Federico, M., Bignami, M., Pichierri, P., and Franchitto, A. (2012). Perturbed replication induced genome wide or at common fragile sites is differently managed in the absence of WRN. Carcinogenesis *33*, 1655–1663.

Naim, V., Wilhelm, T., Debatisse, M., and Rosselli, F. (2013). ERCC1 and MUS81-EME1 promote sister chromatid separation by processing late replication intermediates at common fragile sites during mitosis. Nat. Cell Biol. *15*, 1008–1015.

Paliwal, S., Kanagaraj, R., Sturzenegger, A., Burdova, K., and Janscak, P. (2014). Human RECQ5 helicase promotes repair of DNA double-strand breaks by synthesis-dependent strand annealing. Nucleic Acids Res. *42*, 2380–2390.

Popuri, V., Huang, J., Ramamoorthy, M., Tadokoro, T., Croteau, D.L., and Bohr, V.A. (2013). RECQL5 plays co-operative and complementary roles with WRN syndrome helicase. Nucleic Acids Res. *41*, 881–899.

Richardson, C., Stark, J.M., Ommundsen, M., and Jasin, M. (2004). Rad51 overexpression promotes alternative double-strand break repair pathways and genome instability. Oncogene *23*, 546–553.

Saponaro, M., Kantidakis, T., Mitter, R., Kelly, G.P., Heron, M., Williams, H., Söding, J., Stewart, A., and Svejstrup, J.Q. (2014). RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. Cell *157*, 1037–1049.

Schlacher, K., Christ, N., Siaud, N., Egashira, A., Wu, H., and Jasin, M. (2011). Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. Cell *145*, 529–542.

Schlacher, K., Wu, H., and Jasin, M. (2012). A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. Cancer Cell *22*, 106–116.

Schwendener, S., Raynard, S., Paliwal, S., Cheng, A., Kanagaraj, R., Shevelev, I., Stark, J.M., Sung, P., and Janscak, P. (2010). Physical interaction of RECQ5 helicase with RAD51 facilitates its anti-recombinase activity. J. Biol. Chem. *285*, 15739–15745.

Sofueva, S., Osman, F., Lorenz, A., Steinacher, R., Castagnetti, S., Ledesma, J., and Whitby, M.C. (2011). Ultrafine anaphase bridges, broken DNA and illegitimate recombination induced by a replication fork barrier. Nucleic Acids Res. *39*, 6568–6584.

Svendsen, J.M., Smogorzewska, A., Sowa, M.E., O'Connell, B.C., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2009). Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell *138*, 63–77.

Urban, V., Dobrovolna, J., Hühn, D., Fryzelkova, J., Bartek, J., and Janscak, P. (2016). RECQ5 helicase promotes resolution of conflicts between replication and transcription in human cells. J. Cell Biol. *214*, 401–415.

Wechsler, T., Newman, S., and West, S.C. (2011). Aberrant chromosome morphology in human cells defective for Holliday junction resolution. Nature *471*, 642–646.

Wyatt, H.D., Sarbajna, S., Matos, J., and West, S.C. (2013). Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells. Mol. Cell *52*, 234–247.

Ying, S., Minocherhomji, S., Chan, K.L., Palmai-Pallag, T., Chu, W.K., Wass, T., Mankouri, H.W., Liu, Y., and Hickson, I.D. (2013). MUS81 promotes common fragile site expression. Nat. Cell Biol. *15*, 1001–1007.

Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. Nat. Cell Biol. *16*, 2–9.

CellPress

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-FLAG M2	Sigma-Aldrich	Cat# F1804; RRID: AB_262044
Mouse monoclonal anti-MCM7 Clone DCS-141	Sigma-Aldrich	Cat# M7931; RRID: AB_260666
Mouse monoclonal anti-MUS81 Clone MTA30 2G10/3	Sigma-Aldrich	Cat# M1445; RRID: AB_532259
Rabbit polyclonal anti-phospho-RECQ5 (Ser727)	This paper	N/A
Rabbit polyclonal anti-phospho-Histone H3 (Ser10)	Millipore	Cat# 06-570; RRID: AB_310177
Rabbit polyclonal anti-GFP rabbit polyclonal	Abcam	Cat# ab290; RRID: AB_303395
Goat polyclonal anti-Omni-probe	Santa Cruz	Cat# sc-499-G; RRID: AB_675765
Rabbit polyclonal anti-RECQ5 (full length)	P. Janscak Laboratory	Urban et al., 2016
Rabbit polyclonal anti-RECQ5 (675-991)	P. Janscak Laboratory	Kanagaraj et al., 2010
Rabbit polyclonal anti-RAD51	Santa Cruz	Cat# sc-8349; RRID: AB_2253533
Mouse monoclonal anti-β-Tubulin (D-10)	Santa Cruz	Cat# sc-5274; RRID: AB_2288090
Rabbit polyclonal anti-TFIIH	Santa Cruz	Cat# sc-293; RRID: AB_2262177
Rabbit polyclonal anti-53BP1	Santa Cruz	Cat# sc-22760; RRID: AB_2256326
Mouse monoclonal anti-Cyclin A (B-8)	Santa Cruz	Cat# sc-271682; RRID: AB_10709300
Mouse monoclonal anti-PICH	Millipore	Cat# 04-1540; RRID: AB_11210090
Rabbit polyclonal anti-FANCD2	Novus	Cat# NB100-182; RRID: AB_10002867
Rabbit polyclonal anti-53BP1	Novus	Cat# NB100-304; RRID: AB_10003037
Mouse monoclonal anti-MUS81 Clone MTA30 2G10/3	Abcam	Cat# ab14387; RRID: AB_301167
Anti-FLAG M2 Magnetic Beads	Sigma-Aldrich	Cat# M8823; RRID: AB_2637089
Rabbit IgG control (pre-immune)	P. Janscak Laboratory	N/A
Goat anti-rabbit IgG-HRP	Sigma-Aldrich	Cat# A0545; RRID: AB_257896
Goat anti-mouse IgG-HRP	Sigma-Aldrich	Cat# A4416; RRID: AB_258167
Bovine anti-goat IgG-HRP	Santa Cruz	Cat# sc-2350; RRID: AB_634811
Alexa Fluor 488 Goat Anti-Rabbit IgG (H+L)	Thermo Fisher Scientific	Cat# A-11034; RRID: AB_2576217
Alexa Fluor 594 Goat Anti-Mouse IgG (H+L)	Thermo Fisher Scientific	Cat# A-11005; RRID: AB_2534073
Bacterial and Virus Strains		
Rosetta(DE3)pLysS Competent Cells	Novagen	Cat# 70956
BL21-CodonPlus(DE3)-RIPL Competent Cells	Agilent Technologies	Cat# 230280
Chemicals, Peptides, and Recombinant Proteins		
Isopropyl-beta-D-thiogalactopyranoside (IPTG)	AppliChem	Cat# A1008
Phenylmethanesulfonyl fluoride (PMSF)	AppliChem	Cat# A0999
cOmplete, EDTA-free Protease Inhibitor Cocktail	Sigma-Aldrich	Cat# 11873580001
Lipofectamine RNAiMAX Transfection Reagent	Thermo Fisher Scientific	Cat# 13778150
Creatine Kinase (CK) from rabbit muscle	SERVA Electrophoresis	Cat# 27298
Creatine phosphate disodium salt tetrahydrate	Sigma-Aldrich	Cat# 27920-5g
Doxycycline	Takara Bio	Cat# 631311
Nocodazole	Sigma-Aldrich	Cat# M-1404
KaryoMAX Colcemid Solution in HBSS	Thermo Fisher Scientific	Cat# 15210040
Aphidicolin	Sigma-Aldrich	Cat# A-0781
RO-3306	Sigma-Aldrich	Cat# SML0569
Cytochalasin B	Sigma-Aldrich	Cat# C6762
5-ethynyl-2′-deoxyuridine (EdU)	Thermo Fisher Scientific	Cat# A10044
TransIT-X2 reagent	Mirus	Cat# MIR 6000

(*Continued on next page*)

CellPress

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GenJet in vitro DNA transfection reagent	SignaGen Laboratories	Cat# SL100489-OS
[gamma-P32]Adenosine 5′-triphosphate (3000 Ci/mmol)	Hartman Analytic	Cat# SCP-301
Ni-NTA Agarose	QIAGEN	Cat# 30230
HIS-Select Nickel Affinity Gel	Sigma-Aldrich	Cat# P6611-500ML
Glutathione Sepharose High Performance	GE Healthcare	Cat# 17527902
Chitin Resin	New England Biolabs	Cat# S6651L
SP Sepharose Fast Flow	GE Healthcare	Cat# 17-0729-05
Heparin Sepharose 6 Fast Flow	GE Healthcare	Cat# 17-0998-01
CHT Ceramic Hydroxyapatite	Bio-Rad	Cat# 157-0040
Source 15S	GE Healthcare	Cat# 17-0944-05
Protein A/G PLUS-Agarose	Santa Cruz	Cat# sc-2003
Proteinase K	AppliChem	Cat# A4392,0010
Lambda Protein Phosphatase	New England Biolabs	Cat# P07535S
Synthetic peptide used for preparation of rabbit polyclonal anti-phospho-RECQ5 (Ser727) antibody (sequence: CAHYGGP(pS)PEKKAK)	This study	N/A
CDK1/Cdc2-Cyclin B1	S. Ferrari Laboratory	N/A
Critical Commercial Assays		
QuickChange II Site-Directed Mutagenesis Kit	Agilent Technologies	Cat# 200523
Click-IT Plus EdU Alexa fluor 488 Imaging Kit	Thermo Fisher Scientific	Cat# C10637
ChIP-IT Express Kit	Active Motif	Cat# 53008
QIAquick PCR Purification Kit	QIAGEN	Cat# 28106
LightCycler 480 DNA SYBR Green I Master	Roche	Cat# 04707516001
Deposited Data		
Original data	This study	http://dx.doi.org/10.17632/kvs9dtvdmy.1
Experimental Models: Cell Lines		
U-2 OS	ATCC	ATCC HTB-96
HEK293	ATCC	ATCC CRL1573
HeLa	ATCC	ATCC CCL-2
T-REx-U2OS	Thermo Fisher Scientific	Cat# R712-07
Oligonucleotides		
Oligonucleotides used for preparation of DNA substrates	VBC Biotech	Matulova et al., 2009
FRA3B-forward (sequence: CACTTCCTAACAGGCCCAAA)	Microsynth AG	Lukas et al., 2011
FRA3B-reverse (sequence: CCTCCACTTCTCCTCCCTCT)	Microsynth AG	Lukas et al., 2011
FRA16D-forward (sequence: TCCTGTGGAAGGGATATTTA)	Microsynth AG	Lukas et al., 2011
FRA16D-reverse (sequence: CCCCTCATATTCTGCTTCTA)	Microsynth AG	Lukas et al., 2011
FRA7H-forward (sequence: TAATGCGTCCCCTTGTGACT)	Microsynth AG	Bergoglio et al., 2013
FRA7H-reverse (sequence: GGCAGATTTTAGTCCCTCAGC)	Microsynth AG	Bergoglio et al., 2013
GAPDH-forward (sequence: CCCTCTGGTGGTGGCCCCTT)	Microsynth AG	Lukas et al., 2011
GAPDH-reverse (sequence: GGCGCCCAGACACCCAATCC)	Microsynth AG	Lukas et al., 2011
siCtrl (sequence: CGUACGCGGAAUACUUCGA)	Microsynth AG	Kanagaraj et al., 2010
siRECQ5#1 (sequence: GGAGAGUGCGACCAUGGCU)	Microsynth AG	Kanagaraj et al., 2010
siRECQ5#2 (sequence: CAGGUUUGUCGCCCAUUGGAA)	Microsynth AG	Urban et al., 2016
siMUS81 (sequence: CAGCCCUGGUGGAUCGAUA)	Microsynth AG	Wechsler et al., 2011
Recombinant DNA		
pET21d-MUS81/EME1$_{6xHis}$	Ciccia et al., 2003	N/A
pGex-$_{GST}$MUS81/EME1$_{6xHis}$	Ciccia et al., 2003	N/A
pDEST15-$_{GST}$MUS81	Ciccia et al., 2003	N/A

(*Continued on next page*)

CellPress

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
pTXB1-based vectors for bacterial production of wild-type, K58R, F666A and Δ515-568 forms of RECQ5	Schwendener et al., 2010	N/A
pTXB1-based vectors for bacterial production of RECQ5S727A and RECQ5S727D	This study	N/A
p11d-tRPA (bacterial production of human RPA)	Henricksen et al., 1994	N/A
pET11c-RAD51 K133R (bacterial production of RAD51K133R)	Chi et al., 2006	N/A
pBluescript SK(-)	Addgene	N/A
pAIO-based vector carrying wild-type RECQ5 ORF in frame with 3xFLAG tag	Urban et al., 2016	N/A
pAIO-based vectors carrying K58R, E584D, F666A or S727A forms of RECQ5 ORF in frame with 3xFLAG tag	This study	N/A
pcDNA3.1HisC-based vectors expressing RECQ5 as a fusion with 6xHis-Xpress tag	Schwendener et al., 2010	N/A
pcDNA3.1HisC-based vectors expressing Δ419-515, Δ515-568, Δ569-594, Δ543-607 or Δ571-653 variants of RECQ5 as a fusion with 6xHis-Xpress tag	This study	N/A
pEGFP-C1-based vectors expressing wild-type or Δ515-568 forms of RECQ5 as a fusion with GFP	Kanagaraj et al., 2010	N/A
Software and Algorithms		
MultiGauge V3.2 (used for gel image analysis)	Fujifilm	N/A
ImageJ (used for analysis of immunofluorescence microscopy images)	ImageJ Software	https://imagej.nih.gov/ij
GraphPad Prism 6 for Mac OS X	GraphPad Software	https://www.graphpad.com

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by, the lead contact, Dr. Pavel Janscak (pjanscak@imcr.uzh.ch).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Source of cell lines used in the study is reported in the Key Resources Table.

METHOD DETAILS

Protein expression and purification

MUS81-EME1$_{6xHis}$, $_{GST}$MUS81-EME1$_{6xHis}$ and $_{GST}$MUS81 were expressed in *E. coli Rosetta* (DE3) pLysS from the plasmids pET21d-MUS81/EME1$_{6xHis}$, pGex-$_{GST}$MUS81/EME1$_{6xHis}$ and pDEST15-$_{GST}$MUS81, respectively (Ciccia et al., 2003). Protein production was induced by addition of 0.1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) followed by incubation at 16°C overnight (O/N). Harvested cells (20 g) were resuspended in 80 mL of CBB buffer [50 mM Tris-HCl (pH 7.5), 10% (w/v) sucrose, 2 mM EDTA] containing 200 mM KCl, 0.01% (v/v) NP40, protease inhibitor (PI) cocktail (aprotinin, benzamidine, pepstatin, chymostatin and leupeptin; final concentration of 5 μg/ml each), 1 mM phenylmethanesulfonyl fluoride (PMSF) and 1 mM β-mercaptoethanol (β-ME) (for His-tagged proteins) or 1 mM dithiothreitol (DTT) (for GST-tagged proteins). Cell suspension was sonicated 3 times for 10 s at 80% amplitude in ice bath (Hielscher-Ultrasound Technology sonicator). The lysate was clarified by ultracentrifugation at 100,000 g for 1 hr at 4°C.

For purification of MUS81-EME1$_{6xHis}$, supernatant was loaded on a 20 mL SP Sepharose (GE Healthcare) column. Proteins were eluted with a 200 mL linear gradient of 150-800 mM KCl in buffer K [20 mM K_2HPO_4 (pH 7.5), 10% (v/v) glycerol, 0.5 mM EDTA] supplemented with 0.01% NP40 and 1 mM β-ME. The peak fractions were pooled and incubated with 1 mL of HIS-Select Nickel Affinity Gel (Sigma-Aldrich) for 2 hr at 4°C. The beads were washed with 10 mL of buffer K (NP40, β-ME) containing 150 mM KCl and 10 mM imidazole. The bound proteins were stepwise eluted in 1 mL fractions with buffer K (NP40, β-ME) containing 150 mM KCl and 150, 300, 500 and 1000 mM imidazole, respectively. Peak fractions were loaded onto a 1 mL Heparin Sepharose 6 (GE Healthcare) column and proteins eluted with a 10 mL linear gradient of 150-500 mM KCl in buffer K (NP40, β-ME). The peak fractions were concentrated on Vivaspin Centrifugal Concentrator (30,000 MWCO PES) (Vivaproducts), aliquoted and stored at −80°C.

For purification of $_{GST}$MUS81-EME1$_{6xHis}$ and $_{GST}$MUS81, the clarified cell lysate was incubated O/N at 4°C with 1 mL of Glutathione Sepharose (GE Healthcare) preequilibrated with buffer K (NP40, DTT) containing 350 mM KCl. Bound proteins were eluted with 7 × 1 mL of buffer K (NP40, DTT) containing 350 mM KCl and 150 mM glutathione. Peak fractions were loaded on a 1 mL Heparin Sepharose 6 column and the proteins were eluted with a 10 mL linear gradient of 150-1000 mM KCl in buffer K (NP40, DTT). Peak fractions were concentrated on Vivaspin Centrifugal Concentrator (30,000 MWCO PES), aliquoted and stored at −80°C.

For the production of RECQ5 and its mutants, previously described pTXB1-based vectors were used (Schwendener et al., 2010). Plasmids were transformed into *E. coli* BL21-CodonPlus(DE3)-RIPL strain and protein expression was induced with 0.2 mM IPTG followed by the incubation at 18°C O/N. Cells were pelleted and stored at −80°C. Cell pellets (15 g) were resuspended in 50 mL of buffer CH [20 mM Tris-HCl (pH 8), 500 mM NaCl, 1 mM EDTA, 0.1% (v/v) Triton X-100] supplemented with 0.2 mM PMSF and PI cocktail. After sonication, the cell lysate was clarified by ultracentrifugation at 100,000 g for 1 hr at 4°C. The supernatant was incubated for 2 hr with 3 mL of chitin beads (New England BioLabs) preequilibrated in buffer CH. Beads were transferred to a column and washed with 20 mL of buffer CH. To induce intein cleavage, 4 mL of buffer CH containing 50 mM DTT were added to the beads, and let mix on a rotary shaker O/N at 4°C. Proteins were eluted with buffer T [25 mM Tris-HCl (pH 7.5), 10% (v/v) glycerol, 0.5 mM EDTA] containing 150 mM KCl, 0.01% (v/v) NP40 and 1 mM DTT. Peak fractions were pooled and loaded on a 5 mL CHT Ceramic Hydroxyapatite (Bio-Rad) column. Proteins were eluted with a 30 mL linear gradient of 200-1000 mM KH_2PO_4 in buffer T (NP40, DTT). Peak fractions were pooled and loaded on a 1 mL Heparin Sepharose 6 column equilibrated with buffer T (NP40, DTT) containing 300 mM KCl. The flow-through fraction was applied onto a 1.5 mL Source 15S column (GE Healthcare) and eluted with a 12 mL linear gradient of 300-1000 mM KCl in buffer T (NP40, DTT). Fractions containing RECQ5 were pooled and concentrated on Vivaspin Centrifugal Concentrator (30,000 MWCO PES), aliquoted and stored at −80°C.

RAD51K133R and RPA were purified as described previously (Burkovics et al., 2016; Henricksen et al., 1994).

Nuclease assays

Fluorescein-5,6-isothiocyanate (FITC)-labeled DNA substrates used for nuclease assays were described previously (Matulova et al., 2009). Reactions were carried out at 37°C in ME buffer [25 mM Tris-HCl (pH 7.5), 2.5 mM $MgCl_2$, and 100 mM KCl]. Reaction mixtures, typically 10 μl, contained 4 or 6 nM DNA substrate, 0.2 or 8 nM MUS81-EME1$_{6xHis}$ and indicated concentrations of RECQ5. Where required, ATP was added to a concentration of 2 mM. Reactions were incubated for indicated periods of time and stopped by the addition of SDS (0.1% final) and proteinase K (0.5 mg/ml) followed by further incubation for 5 min. After adding 2 μL of 6x loading dye [10 mM Tris-HCl (pH 7.6), 0.12% (w/v) Orange G, 60% (v/v) glycerol, 60 mM EDTA], the reaction mixtures were separated on 10% (w/v) native polyacrylamide gel in 1x TBE buffer at a constant voltage of 110 V for 1 hr at 4°C. Fluorescent DNA species were visualized by Image Reader FLA-9000 and quantified using MultiGauge V3.2 software (Fujifilm).

For RAD51 inhibition assay, RAD51K133R (500 nM) was preincubated with 3′-flap DNA substrate (6 nM) for 10 min at 37°C in 10 μL of ME buffer supplemented with BSA (50 μg/ml), ATP (2 mM) and ATP-regenerating system (0.1 mg/ml creatine kinase and 1 mM creatine phosphate). Subsequently, reactions were supplemented with indicated concentrations of RECQ5 and incubated for 10 min at 37°C. Finally, MUS81-EME1$_{6xHis}$ (8 nM) was added and incubation was continued for additional 20 min. Reactions were terminated and analyzed as describe above.

D-loop assay

The assay was performed as described previously (Burkovics et al., 2016). Briefly, a fluorescently labeled 90-mer oligonucleotide (3.6 μM nucleotides) was incubated for 5 min with RAD51K133R (1 μM) followed by the addition of RPA (100 nM) and various amounts of RECQ5. Followed by a 4 min incubation, 1 μl HOP2-MND1 (400 nM) was added to a total volume of 10 μl, and the mixture was incubated for an additional 1 min. The reaction was initiated by adding 2 μl of pBluescript SK(-) replicative form I (50 μM base pairs) and incubated for 7 min. Reactions were stopped by addition of SDS (0.05% final) and proteinase K (0.5 mg/ml), incubated at 37°C for 10 min and separated on 0.9% (w/v) agarose gel. Fluorescent DNA species were visualized by Image Reader FLA-9000 and quantified using MultiGauge V3.2 software (Fujifilm).

Gel retardation assay

RAD51K133R was incubated with fluorescently labeled DNA substrate (6 nM) in 10 μL of ME buffer supplemented with BSA (50 μg/ml), ATP (2 mM) and ATP-regenerating system (0.1 mg/ml creatine kinase and 1 mM creatine phosphate) at 37°C for 30 min. Samples were separated on 0.9% (w/v) agarose gel in TAE buffer at a constant voltage of 70 V for 1 hr at 4°C. DNA was visualized by Image Reader FLA-9000.

Plasmid constructions

Derivatives of the pAIO, pEGFP-C1 and pcDNA3.1HisC vectors carrying human RECQ5 ORF in frame with the 3xFLAG tag, GFP and 6xHis-Xpress tag, respectively, were described previously (Urban et al. 2016; Kanagaraj et al., 2010; Schwendener et al., 2010). The RECQ5 ORF in the pAIO vector contains a set of silent mutations conferring resistance of the resulting transcript to siRECQ5#1 (Urban et al. 2016). The vector also encodes for a doxycycline-inducible shRNA equivalent to siRECQ5#1 (Urban et al. 2016). Site-directed mutagenesis of the RECQ5 ORF (K58R, E584D, F666A, S727A and S727D) was performed using a QuickChange II

Site-Directed Mutagenesis Kit (Agilent Technologies). Internal deletions of the RECQ5 ORF in pcDNA3.1HisC were constructed by subcloning from pTXB1-based constructs described previously (Schwendener et al., 2010).

Cell culture

U2OS, HEK293 and HeLa cells were grown in Dulbecco's Modified Eagle Medium (DMEM; Thermo Fisher Scientific) supplemented with 10% fetal calf serum (FCS; Thermo Fisher Scientific) and streptomycin/penicillin (100 U/ml) at 37°C in a humidified incubator containing 5% CO_2. Stable U2OS T-REx cell lines carrying pAIO-based vectors for expression of Flag-tagged versions of wild-type RECQ5, RECQ5K58R, RECQ5E584D, RECQ5F666A, RECQ5S727A and RECQ5Δ515-568, respectively were prepared as previously described (Urban et al., 2016). Cells were cultivated in DMEM supplemented with 10% FBS (Tet-free approved), streptomycin/penicillin (100 U/ml), 50 μg/ml hygromycin B and 1 μg/ml puromycin. To induce expression of RECQ5 variants, doxycycline was added to a concentration of 0.4-1.0 ng/ml. Expression level of recombinant RECQ5 was tuned to be comparable with the level of endogenous RECQ5 by adjusting doxycycline concentration. To synchronize cells in early mitosis, cells were treated with nocodazole (Sigma-Aldrich) at a concentration of 200 ng/ml for 16 hr. After nocodazole treatment, cells were shaken off from the plate, washed two times in PBS and reseeded.

Small-interfering RNA and plasmid transfections

For siRNA transfection, cells were seeded in DMEM complete medium (for U2OS T-REx cells, the medium was supplemented with puromycin and hygromycin) to reach a confluency of 30%–40% at the day of transfection. Transfections of siRNAs (a final concentration of 40 nM) were done using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. Twenty-four hours after siRNA transfection, the medium was exchanged with fresh medium. Treatments with drugs (APH, RO-3306 or nocodazole) were started 48-50 hr after siRNA transfection. siRNA oligoduplexes used in the study were purchased from Microsynth AG. The sequences of the sense strand of siRNA duplexes are: CGUACGCGGAAUACUUCGA (siCtrl, control); GGAGAGUGCGACCAUGGCU (siRECQ5#1); CAGGUUUGUCGCCCAUUGGAA (siRECQ5#2); and CAGCCCUGGUGGAUCGAUA (siMUS81).

For ectopic expression of RECQ5 variants in stable U2OS T-REx cell lines, endogenous RECQ5 was depleted by transfection of siRECQ5 #1 for a total time of 72 hr. Twenty-four hours after siRNA transfection, doxycycline (0.4-1.0 ng/ml) was added to induce expression of RECQ5 variants for a further 48 hr.

Plasmid transfections were done using TransIT-X2 reagent (Mirus) for HEK293 and U2OS T-REx cells, and GenJet reagent (SignaGen Laboratories) for U2OS cells according to the manufacturer's protocols.

Western blotting

Cells were suspended in TN2 buffer [50 mM Tris-HCl buffer (pH 7.5), 120 mM NaCl, 20 mM NaF, 1 mM EDTA, 6 mM EGTA, 15 mM Na-Pyrophosphate and 0.5% (v/v) NP-40] supplemented with 1 mM benzamidine, 0.2 mM PMSF, 0.5 mM sodium orthovanadate and protease inhibitor cocktail (Complete, EDTA-free; Sigma-Aldrich), and sonicated for 5 min with a Diagenode sonicator. Cellular debris was isolated from soluble fraction by centrifugation at 18,000 g for 30 min at 4°C. Samples were subjected to SDS-PAGE. Separated proteins were transferred to a Hybond-P PVDF membrane (GE Healthcare) in a semi-dry apparatus (Sigma) for about 100 min at 0.8 mA/cm^2. The membrane was blocked with 5% non-fat dry milk in TBS-T [20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.1% (v/v) Tween-20] for 1 hr. Afterward, the membranes were incubated with the primary antibodies in 5% milk/TBS-T at 4°C O/N. The membranes were then washed three times in TBS-T and incubated with appropriate horseradish peroxidase-coupled (HRP) secondary antibody for 45 min at RT. Afterward, the membranes were washed three times with TBS-T and bands were detected by luminol-based reaction using a chemiluminescence reagent (Pierce). The primary antibodies used for WB: FLAG mouse monoclonal (F1804, Sigma-Aldrich; 1:500 dilution), MCM7 mouse monoclonal (M7931, Sigma-Aldrich; 1:2000), MUS81 mouse monoclonal (M1445, Sigma-Aldrich; 1:2000), phospho-RECQ5 (Ser727) rabbit polyclonal (Janscak lab; 1:2000), phosho-Histone H3 (Ser10) rabbit polyclonal (06-570, Millipore; 1:1000), GFP rabbit polyclonal (ab290, Abcam; 1:2000), Omni-probe goat polyclonal (sc-499-G, Santa Cruz Biotechnology; 1:1000), RECQ5 rabbit polyclonal (Janscak lab; 1:2000), RAD51 rabbit polyclonal (sc-8349, Santa Cruz; 1:1000), β Tubulin (D-10) mouse monoclonal (sc-5274, Santa Cruz Biotechnology; 1:2000) and TFIIH rabbit polyclonal (sc-293, Santa Cruz Biotechnology; 1:2000). Secondary antibodies used for WB: goat anti-rabbit IgG-HRP (A0545, Sigma-Aldrich; 1:10000), goat anti-mouse IgG-HRP (A4416, Sigma-Aldrich; 1:5000), bovine anti-goat IgG-HRP (sc-2350, Santa Cruz Biotechnology; 1:10000). The anti-phospho-RECQ5 (Ser727) antibody was prepared using a synthetic peptide CAHYGGP(pS)PEKKAK (Clonestar Peptide Services s.r.o.).

Immunoprecipitation

Harvested cells were lysed in TN2 buffer supplemented with 1 mM benzamidine, 0.2 mM PMSF, 0.5 mM sodium orthovanadate and protease inhibitor cocktail (Complete, EDTA-free; Sigma-Aldrich) for 30 min on a rotary shaker at 4°C and sonicated for 5 min with a Diagenode sonicator. The cell lysate was centrifuged at 18,000 g for 30 min at 4°C to obtain whole cell extract. Cell extracts (1 mg of protein) were incubated with 2 μg of affinity-purified rabbit polyclonal anti-RECQ5 antibody (or IgG isolated from a preimmune serum) in a volume of 0.5 - 1 mL of TN2 buffer on a rotary shaker at 4°C O/N. Where indicated, the mixtures were supplemented with ethidium bromide to a final concentration of 50 μg/ml. Next, 20 μl of Protein A/G PLUS-Agarose beads (Santa Cruz) were added and incubated for 2 - 3 hr on a rotary shaker at 4°C. Beads were collected by centrifugation at 2,000 rpm for 3 min at 4°C followed by four washes

with 1 mL of TN2 buffer. Bound proteins were eluted by boiling with 25 μl of Laemmli sample buffer for 5 min at 95°C. Samples were analyzed by western blotting. Extracts of cells expressing Flag-tagged RECQ5 variants (1 mg of protein) were immunoprecipitated using anti-FLAG M2 magnetic beads (10 μl; Sigma-Aldrich). Immunoprecipitates were washed four times with TN2 buffer and analyzed by western blotting.

Pull-down assays

For GST pull-down assays, Glutathione Sepharose (GE Healthcare) was equilibrated with buffer T containing 100 mM KCl, 0.01% (v/v) NP40 and 1 mM DTT. Binding reactions were carried out in a volume of 0.1 mL of buffer T and contained 25 μL of Glutathione Sepharose beads and 5 μg of purified proteins ($_{GST}$MUS81-EME1$_{6xHis}$+RECQ5, $_{GST}$MUS81+RECQ5 or GST+RECQ5). The KCl concentration was adjusted to 100 mM. The mixtures were incubated for 30 min at 4°C with gentle shaking. Samples were centrifuged (6,000 g, 30 s) and the supernatants (unbound fraction) were collected. The beads were then washed three times with 10 volumes of buffer T containing 100 mM KCl and bound proteins were eluted with 20 μL of 2x Laemmli sample buffer. Bound and unbound fractions were analyzed by SDS-PAGE followed by silver or Coomassie Brilliant Blue staining.

For CBD pull-down assays, RECQ5 and its variants fused C-terminally to a chitin-binding domain (CBD) were produced in *E. coli* BL21-CodonPlus(DE3)-RIPL as previously described (Kanagaraj et al., 2010; Schwendener et al., 2010). Cells harvested from a 10 mL culture were resuspended in 1 mL of buffer CH supplemented with 0.2 mM PMSF and a protease inhibitor cocktail (Complete, EDTA-free; Sigma-Aldrich). Cells were disrupted by sonication and the resulting extracts were clarified by centrifugation at 20,000 g for 45 min at 4°C. Cleared extracts (typically 50 μl) were incubated with 25 μL of chitin beads (New England Biolabs) in a total volume of 500 μL of buffer CH supplemented with 0.2 mM PMSF and protease inhibitor mixture for 2 hr at 4°C. The beads were then washed once with 1 mL of CH buffer and three times with 1 mL of buffer TN1 [50 mM Tris-HCl (pH 8), 120 mM NaCl, 0.5% (v/v) Nonidet P-40]. After each wash, beads were collected by centrifugation at 3,000 g for 2 min at 4°C. The washed beads were incubated for 2 hr at 4°C with total extract of HEK293 (600 μg of protein) or 500 ng of purified RAD51 protein in a volume of 500 μL of buffer TN1 supplemented with 0.2 mM PMSF and protease inhibitor mixture, and then washed again three times with buffer TN1. Bound proteins were released from the beads by adding 25 μL of 3x Laemmli sample buffer followed by incubation at 95°C for 5 min. Eluted proteins were analyzed by western blotting.

For Ni-NTA pull-down assays, HEK293 transiently transfected with pcDNA3.1/HisC-based vectors expressing wild-type or mutant forms of RECQ5 as fusions with a 6xHis-Xpress epitope tag were lysed in TN2 buffer supplemented with 1 mM benzamidine, 0.2 mM PMSF, 0.5 mM sodium orthovanadate and protease inhibitor cocktail (Complete, EDTA-free; Sigma-Aldrich). Clarified extract (typically 800 μg of protein) was incubated with 20 μL of Ni-NTA Agarose beads (QIAGEN) for 2 hr at 4°C in a volume of 500 μL of TN2 buffer supplemented with protease/phosphatase inhibitors and 20 mM imidazole. After incubation, the beads were washed three times with 1 mL of TN2 buffer supplemented with 0.2 mM PMSF and 20 mM imidazole. After each washing step, beads were collected by centrifugation at 3,000 g for 2 min at 4°C. Bound proteins were released from nickel beads by adding 25 μL of 3x Laemmli sample buffer followed by incubation at 95°C for 5 min. Eluted proteins were analyzed by western blotting. RECQ5 variants were detected using Omni-probe antibody.

λ-phosphatase assay

To prove that the shift in electrophoretic mobility of RECQ5 induced by nocodazole treatment (200 ng/ml) is due to phosphorylation, cell extracts or RECQ5 immunoprecipitates were treated with lambda protein phosphatase (λ-phosphatase; New England Biolabs). For this, cell lysates were supplemented with benzamidine (1.25 μl/ml) and N-Ethylmaleimide (10 mM). Whole cell extract (40-50 μg of protein) or washed beads (15 μl) from RECQ5 immunoprecipitation (0.5 mg protein) were mixed with 400 U of λ-phosphatase in a 30 μl volume of 1x NEB buffer for PMP (New England Biolabs) supplemented with 1 mM $MnCl_2$. Reactions were incubated at 30°C for 30 min and stopped by adding 5x Laemmli sample buffer followed by incubation at 95°C for 5 min. Samples were analyzed by western blotting along with mock-treated samples.

In vitro phosphorylation assay

Purified RECQ5 protein (1μg) was incubated with 2 U of CDK1/Cdc2-Cyclin B1 (provided by S. Ferrari laboratory), 0.05 mM ATP and 0.1μl [γ- ^{32}P]ATP (3000 Ci/mmol; Hartman Analytic) in a volume of 10 μl of 1x NEB3 buffer for 15 min at 37°C with shaking in a thermomixer (Eppendorf). Reactions were stopped by adding 10 μl of 2x Laemmli sample buffer and heating at 95°C for 5 min. Samples were loaded onto a 8% SDS-PAGE gel and run at 15-20 mA for 2 hr. The gel was stained with Coomassie Brilliant Blue solution and, after destaining, dried on a Whatman paper for 1.5 hr at 80°C and exposed to a phosphor screen (Storage Phosphor Screen, Molecular Dynamics). The screens were scanned with a Typhoon 9400 phosphorimager (GE Healthcare). One unit of CDK1 activity is defined as the amount of enzyme required for the transfer of 1 nmol of phosphate to 1 μg of histone H1 in 1 min at 37°C in a total volume of 10 μl.

Immunofluorescence assays

Cells grown on coverslips, transfected with siRNA or treated with drugs, were fixed with 4% formaldehyde. After several washes with PBS, cells were permeabilized in cold methanol (−20°C) or in 50:50 methanol/acetone mix (−20°C) for 15 min at −20°C. Cells were then washed with PBS and blocked in 5% BSA/PBS solution for 45 min. Coverslips were then incubated O/N at 4°C with appropriate

primary antibody diluted in 5% BSA/PBS. The following day, coverslips were washed several times with PBS and incubated with the secondary antibody for 45 min at RT. After washing with PBS, coverslips were mounted with Vectashield medium containing DAPI. The mounted slides were left to dry at RT for 30 min and then sealed with nail polish. Slides were analyzed with a Leica DM2700 upright fluorescent microscope.

For anaphase bridge analysis (both bulky bridges and UFBs), cells were synchronized with 9 μM RO-3306 (Sigma-Aldrich) for 16 hr before releasing them first in 1x PBS for 5 min at RT and then in DMEM medium for a total time of 1.5 hr at 37°C. After RO-3306 release, cells were simultaneously crosslinked and permeabilized using PTEMF buffer [20 mM PIPES (pH 6.8), 10 mM EGTA, 0.2% (v/v) Triton X-100, 1 mM $MgCl_2$ and 4% (v/v) formaldehyde]. Washing steps were performed in PBS containing 0.2% (v/v) Triton X-100. For UFB analysis, cells were additionally treated with 0.4 μM APH for 24 hr before the release from RO-3306. At least 50 anaphase cells were scored in each experiment.

The primary antibodies used for immunofluorescence staining: 53BP1 rabbit polyclonal (sc-22760, Santa Cruz; 1:200 dilution); Cyclin A mouse monoclonal (sc-271682, Santa Cruz; 1:50); PICH mouse monoclonal (04-1540, Millipore; 1:50), FANCD2 rabbit polyclonal (NB100-182, Novus; 1:1000). The secondary antibodies used for immunofluorescence staining: Alexa Fluor 488 Goat Anti-Rabbit IgG (A11034, Thermo Fisher Scientific; 1:300) and Alexa Fluor 594 Goat Anti-Mouse IgG (A11005, Thermo Fisher Scientific; 1:300).

Preparation and analysis of chromosome spreads

U2OS cells seeded in 10-cm dishes were transfected with indicated siRNAs. Two days post transfection, cells were treated with 0.2 μM APH for 24 hr as indicated in figure legends. Nocodazole (200 ng/ml) was added 5 hr before collection. Cells were harvested by mitotic shake-off and collected in a 15 mL Falcon tube. After centrifugation for 5 min at 4°C (2,000 g), the supernatant was removed and the pellet resuspended in 1 mL of DMEM. Following addition of 8 mL of a 75 mM KCl solution, cells were incubated at 37°C for 15 min. In the meantime, Carnoy's buffer was prepared (75% methanol, 25% glacial acetic acid) and 5 mL of this buffer were added to the cell suspension for 15 min at RT. The fixing steps with Carnoy's buffer were repeated 3 times with 8 mL of Carnoy's buffer added for each incubation. Cells were then spread on slides drop-wise and let dry O/N at RT. The following day, dried slides were mounted with Vectashield medium containing DAPI. Chromosome spreads were analyzed on a Leica DM2700 upright fluorescent microscope and visible chromatid breaks and gaps were counted. At least 50 chromosome spreads were scored in each experiment.

EdU labeling and detection in mitotic cells

Asynchronously growing U2OS cells were synchronized in late G2 phase of the cell cycle by incubation with 9 μM RO-3306 for 16 h along with 0.4 μM APH. Cells were subsequently washed three times with PBS for 5 min and then released into fresh medium (pre-warmed to 37°C) containing 20 μM 5-ethynyl-2′-deoxyuridine (EdU) and 0.1 μg/ml Colcemid (Thermo Fisher Scientific) for up to 60 min. Cells were then harvested, pelleted at 1,000 rpm for 5 min, and then swollen by incubation in 75 mM KCl (pre-warmed to 37°C) for 20 min at 37°C. Swollen mitotic cells were collected at 1,000 rpm for 5 min, fixed using methanol:acetic acid (3:1), and dropped onto pre-hydrated glass slides (Thermo Fisher Scientific) and aged for up to 24 hr followed by EdU detection using Click-IT Plus EdU Alexa fluor 488 Imaging Kit (Thermo Fisher Scientific). Chromosomes were stained using DAPI (Vectashield; Vector Labs). Images were captured using a Leica SP8 upright confocal laser-scanning microscope (63x/1.40 OIL) objective.

For measurement of the frequency of FANCD2 foci co-localizing with EdU foci in intact prometaphase nuclei, cells were released from G2 arrest for 35 min in medium containing 20 μM EdU. Mitotic cells were shaken off and re-seeded in poly-lysine-coated glass slides (Sigma-Aldrich) followed by simultaneous fixation and permeabilization using PTEMF buffer for 20 min at RT. EdU was detected using Click-IT chemistry prior to immunofluorescence staining of FANCD2. Images were captured using an Olympus BX63 microscope, and analyzed using ImageJ. FANCD2 was detected using the NB100-182 anti-FANCD2 antibody (Novus) at a dilution of 1:1000.

Analysis of micronuclei

Cells were grown on coverslips in 6-well plates. Cell culture medium was supplemented 16 hr before fixation with 2 μg/ml cytochalasin B (Sigma-Aldrich) to block cells in cytokinesis. Cells were fixed with 4% formaldehyde for 10 min and mounted with Vectashield mounting medium containing DAPI. Slides were analyzed with a Leica DM2700 upright fluorescent microscope. For quantification, only DAPI-stained binucleated cells were counted, and distinct micronuclei in the vicinity of these cells were considered as positive. At least 150 binucleated cells were scored in each experiment.

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) experiments were done with the ChIP-IT Express Kit (Active Motif) as described previously with minor modifications (Kanagaraj et al., 2010). U2OS cells were seeded in a 10-cm plate and transfections of indicated siRNAs were performed twice at 24 hr and 48 hr post seeding. Transient transfections of constructs expressing GFP-tagged RECQ5 variants, or empty vector were performed 24 hr post seeding. Where indicated, APH (0.4 μM) and/or RO-3306 (9 μM) were added to cells two days after first siRNA or plasmid transfection. Cells were crosslinked 22 hr later with 1x formaldehyde solution [1.1% formaldehyde, 10 mM NaCl, 0.1 mM EDTA (pH 8.0), 5 mM HEPES (pH 7.9)] at RT for exactly 15 min, followed by addition of glycine solution (125 mM glycine) and incubation for 5 min at RT to quench the crosslinking reaction. Cells were harvested and collected by

centrifugation in a refrigerated centrifuge at 1,200 rpm for 10 min. Cells were then washed twice with PBS containing 0.5% Igepal. Chromatin fragments used in immunoprecipitation reactions were prepared by shearing of crosslinked chromatin using a bioruptor sonication device (Diagenode). Ten percent of the sonicated chromatin was stored at −80°C for use as an input sample. For each ChIP reaction, sonicated chromatin (7.5 μg) was immunoprecipitated O/N at 4°C with either indicated antibody or a control IgG (4 - 5 μg each). After washing, immunocomplexes were eluted from the beads and de-crosslinked according to the manufacturer's instructions (Active Motif). ChIP and input samples were purified with the QIAquick PCR Purification Kit (QIAGEN) and DNA was eluted with 50 μL of sterile water. At least three independent experiments were performed for each condition. In each experiment, eluted DNA samples were subjected in triplicates (2-3 μL each) to quantitative real-time PCR (qPCR) analysis on a Roche LightCycler 480 Real-time PCR system using LightCycler 480 DNA SYBR Green I Master (Roche). Data were analyzed using the Pfaffl's method as described previously (Kanagaraj et al., 2010). Fold enrichment of each target region was calculated as ratio of the amounts of immunoprecipitated DNA estimated for the specific antibody versus control IgG. For plasmid transfection experiments, fold enrichment was calculated as a ratio of the qPCR values obtained with RECQ5-transfected cells versus mock (EV)-transfected cells. Antibodies used for ChIP in this study: 53BP1 rabbit polyclonal (NB100-304, Novus Biologicals), MUS81 mouse monoclonal (ab14387, Abcam), RAD51 rabbit polyclonal (sc-8349, Santa Cruz), RECQ5 rabbit polyclonal (Janscak lab) and GFP rabbit polyclonal (ab290, Abcam). The sequences of primers used for qPCR are: CACTTCCTAACAGGCCCAAA (FRA3B-forward); CCTCCACTT CTCCTCCCTCT (FRA3B-reverse); TCCTGTGGAAGGGATATTTA (FRA16D-forward); CCCCTCATATTCTGCTTCTA (FRA16D-reverse); TAATGCGTCCCCTTGTGACT (FRA7H-forward); GGCAGATTTTAGTCCCTCAGC (FRA7H-reverse); (CCCTCTGGTGGTG GCCCCTT (GAPDH-forward); and GGCGCCCAGACACCCAATCC (GAPDH-reverse).

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical significance of differences between datasets from EdU incorporation experiments was determined by Mann-Whitney U test using GraphPad Prism 6 software. Statistical details of experiments can be found in the figure legends.

DATA AND SOFTWARE AVAILABILITY

Original imaging data and data from chromosome spread analyses have been deposited to Mendeley Data and are available at http://dx.doi.org/10.17632/kvs9dtvdmy.1.

7.2 Efficient mRNA processing prevents gene gating-associated replication stress in mammalian cells

Federico Teloni[1], Jone Michelena[1], Sinan Kilic[1], Jana Dobrovolna[3], Ralph Imhof[1], **Shruti Menon[2]**, Christina Ambrosi[1], Tuncay Baubec[1], Pavel Janscak[2,3], Matthias Altmeyer[1,*]

[1] Department of Molecular Mechanisms of Disease, University of Zurich, CH-8057, Zurich, Switzerland

[2] Institute of Molecular Cancer Research, University of Zurich, CH-8057, Zurich, Switzerland

[3] Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Videnska 1083, 143 00

[*] Correspondence (lead contact): Matthias Altmeyer, matthias.altmeyer@uzh.ch

Manuscript submitted

I participitated in this study by contributing to the generation of the U2OS T-REx stable cell line expressing catalytically-inactive RNase H1[RNH1(D210N)-GFP], used for the detection of R-loops. I also performed a set of DNA fiber experiments.

8. CURRICULAM VITAE

Personal Details

Name:	**Shruti**	Date of birth:	**09.11.1988**
Surname:	**Menon**	Nationality:	**Indian**

Education and training

2013-2018	**Ph.D. in Cancer Biology**, IMCR, University of Zurich, Switzerland Supervisor: PD Dr. Pavel Janscak Project title: Role of co-transcriptional R-loops in DNA replication stress.
2011-2013	**Research assistant,** Indian Institute of Science, Bangalore, India Supervisor: Prof. Valakunja Nagaraja Project title: Studying the enzyme Topoisomerase I from Mycobacterium tuberculosis and validating its essentiality by generation of conditional knockdown systems.
2009-2011	**Masters in Biotechnology**, Vellore Institute of Technology, Vellore, India **Master thesis:** School of Biotechnology, Jawaharlal Nehru University, New Delhi, India Supervisor: Prof. Aparna Dixit
2006-2009	**Bachelors in Biochemistry**, Sri Venkateswara College, University of Delhi, New Delhi, India

Fellowship and awards

2017	Cancer Biology PhD Program Travel award to attend "Eukaryotic DNA Replication & Genome Maintenance Meeting", CSHL, USA
2015	The University Research Priority Program (URPP) grant for PhD students,University of Zurich

Publications

Di Marco, S., Z. Hasanova, R. Kanagaraj, N. Chappidi, V. Altmannova, **S. Menon**, H. Sedlackova, J. Langhoff, K. Surendranath, D. Huhn, R. Bhowmick, V. Marini, S. Ferrari, I. D. Hickson, L. Krejci and P. Janscak (2017). "RECQ5 Helicase Cooperates with MUS81 Endonuclease in Processing Stalled Replication Forks at Common Fragile Sites during Mitosis." Mol Cell **66**(5): 658-671 e658.

Ahmed, W., A. G. Bhat, M. N. Leelaram, **S. Menon** and V. Nagaraja (2013). "Carboxyl terminal domain basic amino acids of mycobacterial topoisomerase I bind DNA to promote strand passage." Nucleic Acids Res **41**(15): 7462-7471.

Ahmed, W., **S. Menon**, A. A. Godbole, P. V. Karthik and V. Nagaraja (2014). "Conditional silencing of topoisomerase I gene of Mycobacterium tuberculosis validates its essentiality for cell survival." FEMS Microbiol Lett **353**(2): 116-123.

Ahmed, W., **S. Menon**, P. V. Karthik and V. Nagaraja (2015). "Reduction in DNA topoisomerase I level affects growth, phenotype and nucleoid architecture of Mycobacterium smegmatis." Microbiology **161**(Pt 2): 341-353.

Ahmed, W., **S. Menon**, D. N. B. K. PV and V. Nagaraja (2015). "Autoregulation of topoisomerase I expression by supercoiling sensitive transcription." Nucleic Acids Res.

9. ACKNOWLEDGEMENTS

First and foremost, I would like to thank PD.Dr. Pavel Janscak for giving me the opportunity to work in his laboratory. Thank you for always being available for discussions and giving us the freedom to think and work independently. It has been a great learning expirience working with you.

I would like to thank my thesis committee, Prof. Kerstin Gari,Prof. Alessandro Sartori, Prof. Primo Schär; for their valuable inputs during the course of my PhD.I would like to express my sincere gratitude to Dr. Jana Dobrovolna, for laying the foundation on which this project stands. Thank you for your help and support throughout these 4 years.

A major challenge of this project was high-throughput sequencing and analysis of the data. For this, I am thankful to Catharine Aquino and Lennart Opitz at the Functional Genomics Center, University of Zurich. I would also like to acknowledge Elriedy Amr Abduo Elsayed for his efforts to help out with the data analysis. Importantly, a heartfelt thank you to Dr. Michal Kolar, for coming to our rescue by providing the much needed bioinformatic support.

I would like to thank all the past members of the Janscak lab – Andreas Sturzenegger and Stefano Di Marco for introducing me to the lab during the start of my PhD; Adisa Hadzic and Domino Schlegal for becoming wonderful friends in the short span of time they were here. A big thank you to the current members of the lab- Nagaraja Chappidi for his scientific inputs; Anca Irina Mihai and Esin Isik for the amazing work environment and for our outside lab adventures. Special thanks to Christiane König for her technical assistance and support, and also for being a good friend. I would like to thank all the members of IMCR for their help, useful discussions and valuable inputs.

My stay in Zurich could not have been this enjoyable without some amazing people - Wareed, Ashwini, Archana and Stefano. I want to specially thank Lepakshi

and Roopesh, for being not just aweome friends but also my family here. A big thank you to Veena, we go such a long way, for patiently reading my thesis, giving me very valuable inputs and for motivating me at every step.

Last but not the least, I want to acknowledge my wonderful family for their constant support and encouragement. A special thank you to my little sister Preeti, for being my best friend and support system through all these years.